Reaction Mechanisms of Collision and Electron Induced Peptide Dissociation Revealed by Ion Spectroscopy

Lisanne Kempkes
Reaction Mechanisms of Collision and Electron Induced Peptide Dissociation
Revealed by Ion Spectroscopy
PhD Thesis, Radboud University, Nijmegen, The Netherlands
Cover design by Rosalie Kempkes
Printed by Ipskamp Printing

ISBN: 978-94-028-1379-1

Reaction Mechanisms of Collision and Electron Induced Peptide Dissociation Revealed by Ion Spectroscopy

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op
dinsdag 19 maart 2019
om 16:30 uur precies

door

Liesbeth Johanna Maria Kempkes

geboren op 4 december 1990
te Wijchen

Promotor:
Prof. dr. J. Oomens

Copromotor:
Dr. J.K. Martens

Manuscriptcommissie:
Prof. dr. F.P.J.T. Rutjes (voorzitter)
Prof. dr. G.J. von Helden
Dr. R. Spezia (Sorbonne Université, Parijs, Frankrijk)

.

Dit werk is mede mogelijk gemaakt door financiering van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), Chemische Wetenschappen, middels VICI-project nr. 724.011.002. Berekeningen in dit werk zijn uitgevoerd op de Nederlandse nationale e-infrastructuur met steun van Coöperatie SurfSARA en het NWO-programma Rekentijd. Het experimentele deel van dit werk is uitgevoerd binnen het FELIX Laboratorium.

Table of Contents

Part I – General introduction and methods

Chapter 1

Introduction

Proteins are produced by living organisms and can have all kinds of functions and appearances. Protein functions include transport, catalysis of metabolic reactions, DNA replication and translation. Enzymes, antibodies and hormones are all examples of proteins. The function of a protein is closely related to its structure. In order to understand the function of a protein, its structure can be studied. Several methods have been developed to study the protein structure, each with its own advantages and disadvantages. This chapter focusses on determining protein structure by using Mass Spectrometry (MS).

Protein structure determination and identification is based on cleavage of the protein in (sequence) fragments. The primary structure of a protein is derived from the differences in the mass to charge ratio between the sequence fragments. Protein identification is in most cases achieved by comparing protein fragmentation patterns with database information, although *de novo* sequencing is also possible. However, this is often not as simple as it may seem, and algorithms are continuously being updated to improve the matching scores.

To date, despite its widespread use in protein sequencing, detailed chemical pathways of the peptide fragmentation reactions occurring in the mass spectrometer are only partially understood. Originally, it was believed that these fragments are simply truncated peptides with linear structures as a consequence of direct cleavage of chemical bonds in the protein backbone. However, detailed MS/MS studies as well as theoretical investigations suggested that the backbone cleavages are often accompanied by rearrangement reactions, leading to fragments possessing cyclic structures. The large variety of chemical moieties and nucleophiles present within peptides, leads to a rich, but complex, fragmentation chemistry, specifically determined for each peptide by the sophisticated interplay between thermodynamic and kinetic effects. The overall aim of the studies presented in this thesis is to obtain a better understanding of the chemical reaction mechanisms underlying the fragmentation process. A combination of IR laser spectroscopy and tandem mass spectrometry enables us to reveal the molecular structures of the peptide fragments, which can eventually improve mass spectrometry based protein characterization.

1.1. Amino acids, peptides and proteins

Although the overall structure of proteins can be very complex, the primary structure of a protein is a simple sequence of amino acid residues, which are the building blocks of any protein. The simplest amino acid structure, glycine, is presented in Figure 1-1. It consists of an NH_2-group (N-terminus), a COOH group (C-terminus), and a side chain (R) which is an hydrogen atom for glycine, all attached to the central α-carbon atom (C_α).[1,2] Additional carbons in the side chain of the amino acids other than glycine are designated by β (2nd position), γ (3rd position) and so forth.

There are 20 naturally occurring amino acids listed in Figure 1-2, each possessing a unique side chain with its own chemical properties such as structure, size, basicity/acidity and solubility. The full names of the amino acids are abbreviated by a three-letter or single-letter code, which are listed in Table 1-1.

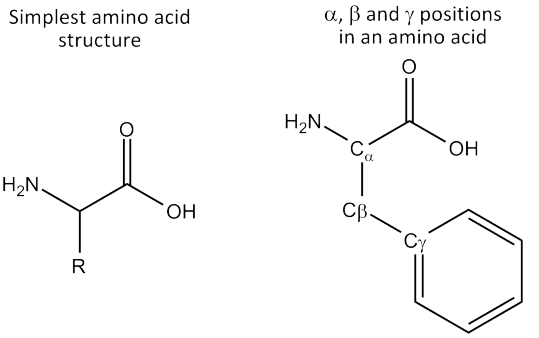


Figure 1-1. *The left panel shows the most simple amino acid structure, the right panel shows the α,β, and γ positions of carbon atoms in an amino acid molecule.*

Amino Acid	Abbreviation		Amino Acid	Abbreviation	
	3- letter	1-letter		3- letter	1-letter
Glycine	Gly	G	Glutamic Acid	Glu	E
Alanine	Ala	A	Asparagine	Asn	N
Valine	Val	V	Glutamine	Gln	Q
Iso-leucine	Ile	I	Threonine	Thr	T
Leucine	Leu	L	Serine	Ser	S
Arginine	Arg	R	Phenylalanine	Phe	F
Lysine	Lys	K	Tyrosine	Tyr	Y
Cysteine	Cys	C	Tryptophan	Trp	W
Methionine	Met	M	Histidine	His	H
Aspartic Acid	Asp	D	Proline	Pro	P

Table 1-1. *List of 3- and 1-letter abbreviations of all 20 in nature occurring amino acids.*

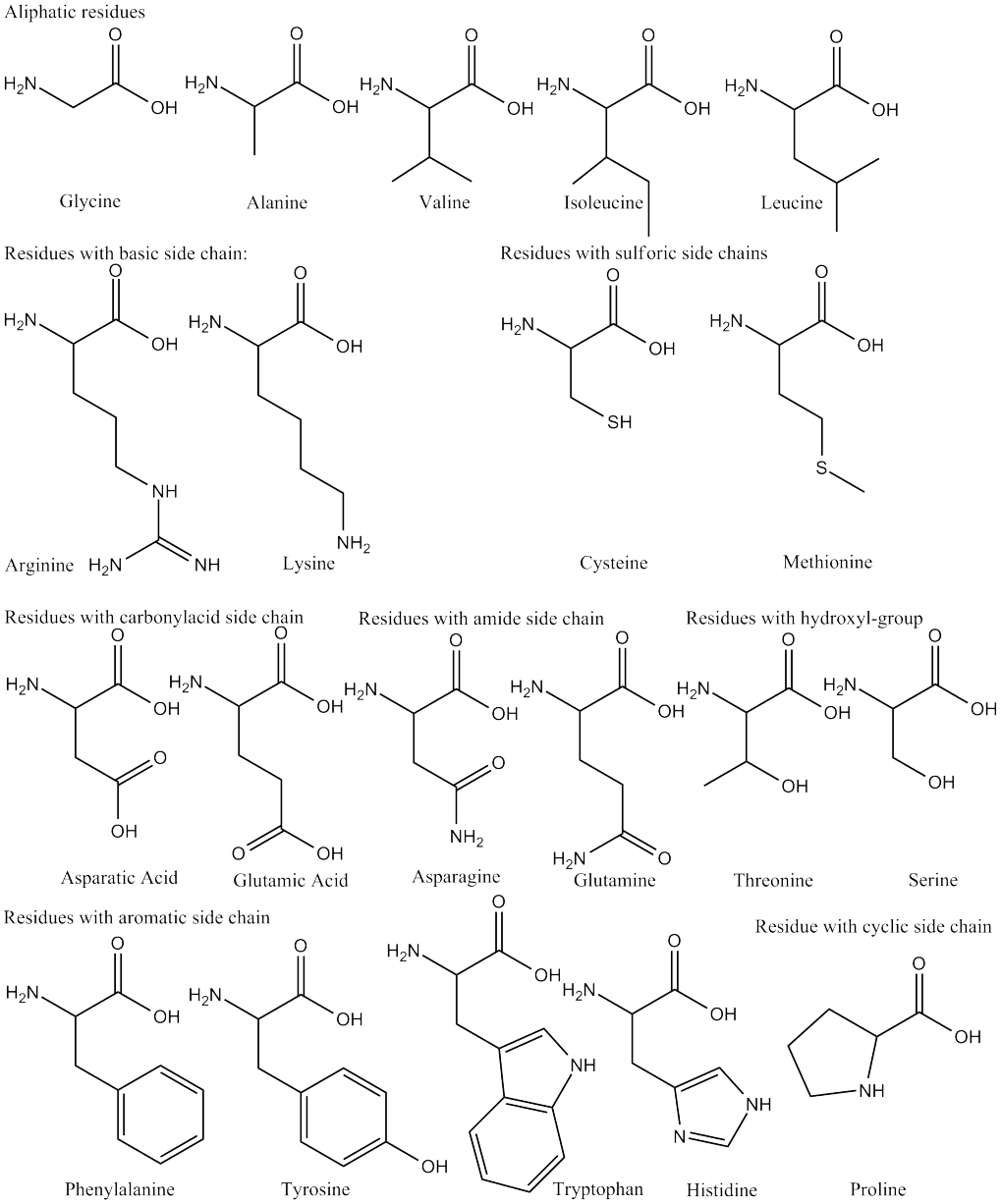


Figure 1-2. *The structures of all 20 natural occurring amino acids, classified on their side chain properties.*

Amino acids are covalently linked to each other to form peptides and proteins by peptide bonds, which are formed between the carbonyl group of the first amino acid and the amino group of the second amino acid upon expulsion of a water molecule (H_2O). This condensation reaction takes place in a ribosome and is illustrated in Figure 1-3. Protein and peptide sequences are by convention indicated with the amino N-terminal end on the left and the carboxylic acid C-terminus is on the right. Two joined amino acids are referred to as a dipeptide, three joined amino acids as a tripeptide, and so on.

Although 20 unique amino acid building blocks may appear to be a small number, a huge diversity of peptides and proteins can be formed. For example a tripeptide can have 8000 combinations. Moreover an identical composition of residues, can have many different sequences. A peptide that consists of more than ~10 amino acid residues is called a polypeptide. A protein is a polypeptide possessing a large number of amino acid residues, the length of a protein can vary from dozens to a complex with millions of amino acids. By joining the 20 different amino acids in different combinations or sequences, proteins with strikingly different properties, functions and activities can be formed.[1,2]

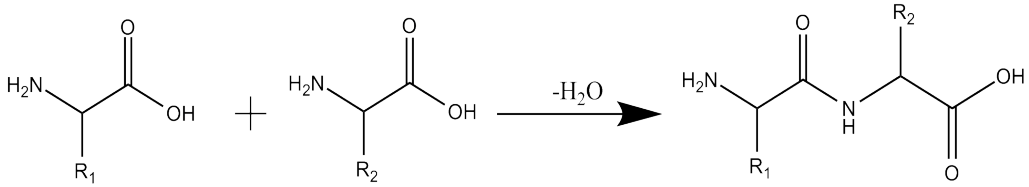


Figure 1-3. *Formation of a dipeptide from two amino acids by an enzymatic condensation reaction.*

1.2. Protein structure and proteomics

Sequence directs structure, and the three-dimensional structure of two molecules with similar sequences can be assumed to be very similar. The amino acid sequence is the primary structure of a protein. Specific hydrogen-bond interaction motifs within a protein are referred to as the secondary structure of a protein. Examples of the secondary structures are α-helices and β-sheets. The tertiary structure of a protein is the overall three-dimensional conformation, e.g. compact or non-compact conformations. The noncovalent association of polypeptides to form a multimeric complex defines the quaternary structure of a protein.[1,2] Figure 1-4 illustrates the primary structure of a protein, its folding into α-helices and β-sheets, the tertiary structure and complex quaternary structure of a protein.

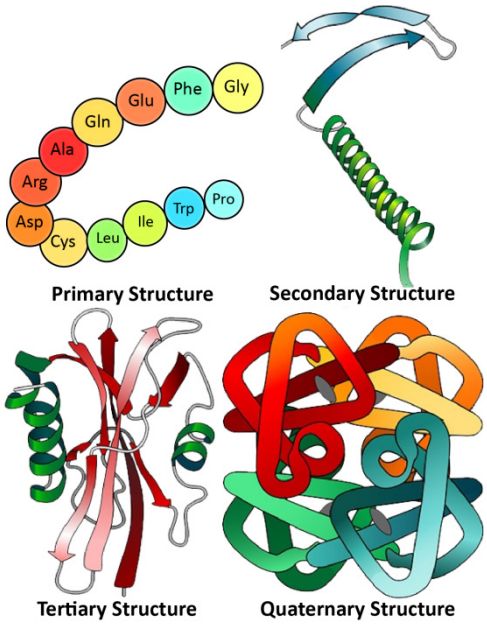


Figure 1-4. *An illustration of the primary, secondary, tertiary and quaternary structure of a protein.*

The proteome is the entire set of proteins that are produced or modified by an organism and is studied in proteomics. The word proteome is derived from PROTEins Expressed by a genOME, which is the complete protein population of a living organism. Studying the proteome of organisms is more complex than studying the genome as the proteome differs from cell to cell and from time to time.[1]

Deoxyribonucleic acid (DNA) contains the genetic information in every organism. DNA is constructed from four nucleotides: adenine (A), cytosine (C), guanine (G) and thymine (T). From DNA, ribonucleic acid (RNA) is encoded, involving the same nucleobases as in DNA, except for thymine which is replaced by uracil (U) in RNA. This process of transcription thus copies the DNA code into an RNA sequence. From the RNA code, the protein sequence is synthesized in a process called translation.[3] Figure 1-5 presents a schematic overview of the flow of genetic information within a biological system to form proteins, according to the central dogma of molecular biology. Three consecutive RNA bases code form one amino acid as shown in Figure 1-5. The primary protein structure is thus defined by the genetic code, while hydrogen-bonding and other non-covalent intramolecular interactions within the protein define the three-dimensional, secondary and tertiary structures.

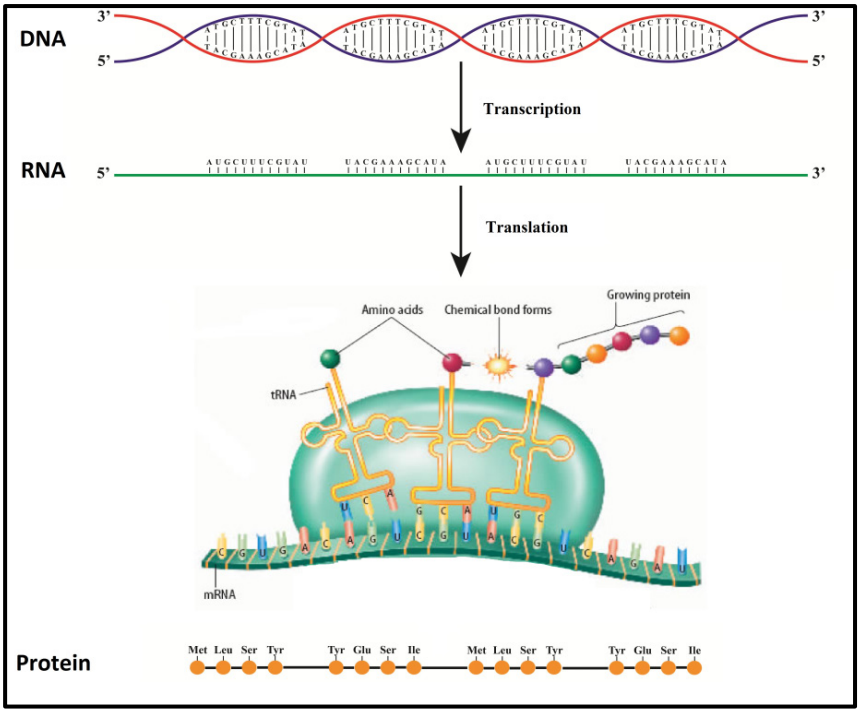


Figure 1-5. *Overview of protein formation from the DNA, DNA codes for RNA by transcription. Proteins are formed by translation of the RNA. Figure adapted from:* [4,5].

A protein is a dynamic complex as the three-dimensional structure changes over time. After the formation of a protein via translation from RNA, proteins are not in their "final" state. For example, new hydrogen bonds can be formed and existing ones can be broken. Moreover, covalent modifications can alter proteins after translation, for instance in phosphorylation, glycosylation, hydroxylation, or methylation reactions, but also by attachment of prosthetic groups, formation of disulfide bonds, deamidation (loss of NH_3) and dehydration (loss of H_2O). These modifications are referred to as Post-Translational Modifications (PTMs), which partly occur spontaneous in the protein or are catalyzed by enzymes. PTMs play a role in misfolding, misassembly or malfunctioning of proteins, which can lead to numerous diseases.[6-10] An example is congenital adrenal hyperplasia, wherein an enzyme that is responsible for the synthesis of cortisol has a hydroxylase deficiency leading to defects in the hormone regulation and gender development in unborn children.[11] Besides this example, PTMs are believed to play a role in aging and aging related-diseases.[12-16]

1.3. Methods for characterization of protein structure

The three-dimensional structure of a protein can be determined by various techniques, all with their own advantages and disadvantages. A prominent example is X-ray crystallography, in which protein crystals are studied by X-ray diffraction.[17,18] This technique gives detailed protein structures, but the protein needs to be crystallized and analysis of the resulting diffraction patterns is challenging. Insoluble proteins, such as membrane proteins are notoriously difficult to crystallize and thus to study by X-ray diffraction. Another often used technique is Nuclear Magnetic Resonance (NMR), which has as an advantage that it can be used on peptides in solution.[19,20] Furthermore, the dynamics of a protein can be followed, such as conformational changes, folding and interaction with other molecules. However, this technique only works for smaller proteins (<20 kDa). Other methods which are used for protein structure determination are among others, cryo-electron microscopy (cryo-EM)[21-23], electron spin resonance (ESR) spectroscopy[24], electron paramagnetic resonance (EPR) spectroscopy[25], and Raman spectroscopy.[1,2,26]

1.3.1. Determining the amino acid sequence

The amino acid sequence of a protein can be determined using several methods, which are briefly summarized here.

Direct from DNA

Protein sequences can be derived indirectly from the DNA sequences in the genome database, since DNA encodes the amino acid sequence of proteins. However, modifications that occur after proteins are bio-synthesized (PTMs), are not stored in the DNA code.[1,2]

Edman degradation

To sequence an entire polypeptide, a chemical method devised by Pehr Edman[27] was commonly employed until the development of MS-based methods. The Edman degradation procedure labels and removes the amino-terminal residue from a peptide, leaving all other peptide bonds intact. Each reaction with the amino-terminal residue can go essentially to completion without affecting any of the other peptide bonds in the peptide. After removal and identification of the amino-terminal residue, the new amino-terminal residue exposed can be labeled, removed and identified through the same series of reactions. This procedure is repeated until the entire sequence is determined. The Edman degradation is

carried out in a machine, known as a sequenator, that mixes reagents in the proper proportions, separates the products, identifies them, and records the results. Often, the complete amino acid sequence can be determined starting with only a few micrograms of protein. However, at each cycle, peptides that did not react in earlier cycles would contribute amino acids to an ever-increasing background, eventually making it impossible to determine which amino acid is next in the original peptide sequence.[1,2]

Sequencing approaches using mass spectrometry

Since the introduction of soft-ionization techniques to ionize and transfer intact proteins to the gas-phase, mass spectrometry has become a common method for protein sequencing. Electrospray ionization (ESI)[28] (see chapter 2 for a more detailed explanation) and matrix-assisted laser desorption ionization (MALDI)[29] are examples of techniques which made it possible to bring the fragile, large and non-volatile biological molecules into the mass spectrometer. In mass spectrometry, the mass-to-charge ratio (m/z) of an injected sample is measured. Although the masses of proteins may not be unique, the sequence of proteins is.

Within mass spectrometry based proteomics, a distinction is made between top-down, middle-down and bottom-up sequencing approaches.[30] In bottom-up proteomics, a protein is cleaved into peptides of on average 10 residues prior to introduction into the mass spectrometer. The most common way to cleave a protein is using an enzyme (protease) that cleaves the peptide backbone after a known amino acid residue. An often used protease is trypsin, which cleaves at the C-terminus of basic amino acid residues, Arg and Lys, resulting in tryptic peptides.[31,32] In middle-down proteomics enzymes with alternative cleavage sites are used generating longer peptides.[31,33] Obviously, the advantage is that the sample to be sequenced is less complex. In top-down proteomics large, intact proteins are fragmented and analyzed[31,34] to reveal their sequence. This generally requires more sophisticated MS platforms and has the advantage that no proteolytic digestion of the proteins is needed.[34]

In tandem-mass spectrometry (MS/MS), a molecule is isolated by m/z value and then dissociated typically by collisional activation. The resulting fragments are analyzed. MS^n involves multiple stages of mass isolation and fragmentation with *n-1* indicating the number of fragmentation steps. The identification of a particular protein does not necessarily require a complete sequence identification as only a very minor portion of all possible protein sequences is realized in nature.[35] The

fragmentation pattern alone usually provides a unique fingerprint that can be associated with a unique protein sequence through a database search for fragment analysis.[35] Fragmentation spectra are compared with in silico databases such as MASCOT[36], SEQUEST[37], TANDEM[38] and OMSSA[39]. A search through the database will often not lead to an exact match, but to a probability score for a match, and the possibility remains that single amino acid differences occur.

In *de novo* peptide sequencing, the amino acid sequence is determined by tandem mass spectrometry[30,35,40-42] without use of a database. *De novo* sequencing is used for proteins that do not give a satisfactory match in database searches. The assignment of an amino acid sequence is obtained by relating back the masses of the fragments to the original peptide structure as the protein is dissociated in fragments that differ by one amino acid residue. The resulting fragments are called sequence ions. Examples of *de novo* sequencing programs are PEAKS[43], PepNovo[44], NovoHMM[45], MSnovo[46] and Vonode[47].

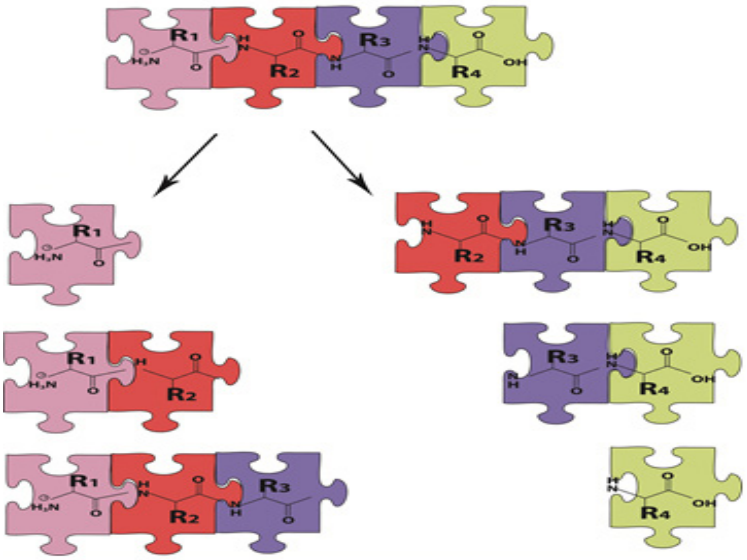


Figure 1-6. *The fragmentation of the peptide backbone illustrated by pieces of a puzzle.*

Figure 1-6 presents the dissociation of a peptide at the various peptide bonds, leading to puzzle pieces that have a unique mass. As the masses of all 20 naturally occurring amino acids are different (except for Leu and Ile), the amino acid residue assignment is based on the mass difference between the sequence ions in the MS/MS spectrum. A mass spectrum that is analyzed by *de novo* peptide sequencing

is presented in Figure 1-7. The sequence ions in this figure are indicated as b- and y-peaks and the mass differences between the series of b-ions and/or in the series of y-ions lead to the assignment of the sequence. In automated *de novo* sequencing a computer algorithm searches for differences between mass peaks that correspond to masses of one of the 20 naturally occurring amino acid residues.[41,46,48-52]

In addition to the sequence ions, many peaks are visible in the mass spectrum in Figure 1-7, some of them having significant intensities. These unidentified peaks (non-sequence ions) are usually not taken into account in the sequence determination and are due to cleavages at sites different from the backbone of the peptide, for example by loss of neutral molecules or side chains. However, these peaks may contain important information about the original protein structure, and can be useful for improved peptide identification, sequence analysis and to detect PTM's.[1,2,53]

The key to include these unidentified peaks in an improved sequence analysis is to better understand the fragmentation reactions and the resulting non-sequence ions in terms of mass but especially also in terms of molecular structure. Using the additional information from non-sequence ions is expected to lead to better probability scores and thus to more reliable protein sequences and structures. The various techniques to fragment a protein that are currently being used in MS-based protein sequencing are outlined in Sections 1.5 and 1.6. The technique used for molecular structure determination of peaks in the mass spectrum is central to this thesis and is explained in Section 1.8.

1.4. Peptide fragmentation

Mass spectrometry has rapidly evolved to become the platform of choice for proteomic analysis. Different fragmentation techniques have been developed, with differences in ion activation, for example: Collision Induced Dissociation (CID[54-57]), Electron Capture Dissociation (ECD[58]), Electron Transfer Dissociation (ETD[59]), Infrared Multiple Photon Dissociation (IRMPD[60]), Blackbody Infrared Radiative Dissociation (BIRD[61]) and Ultraviolet Photodissociation (UVPD[62]).[30]

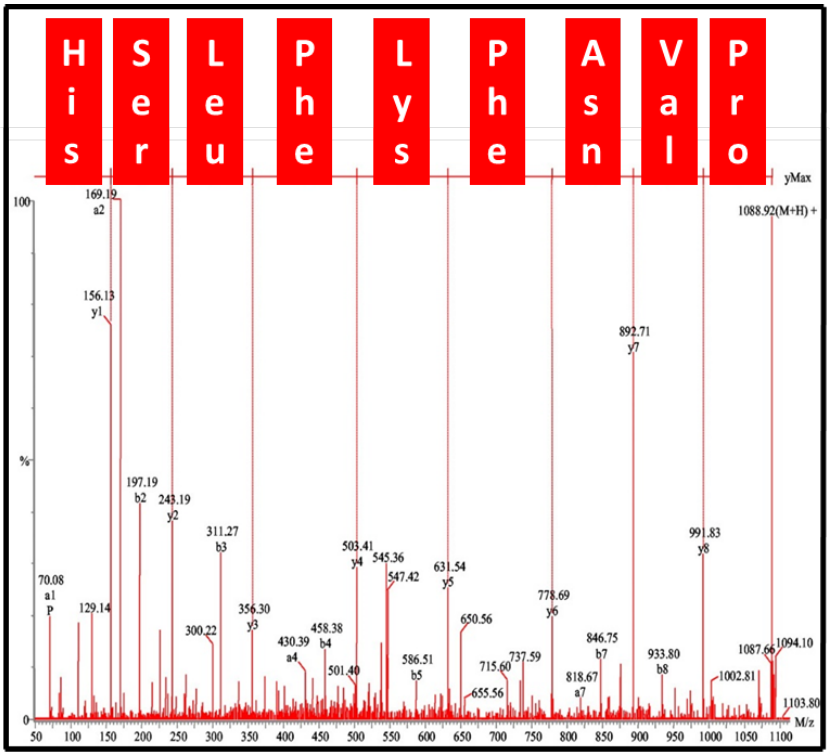


Figure 1-7. *Example of a mass spectrum used for peptide sequencing, based on the mass difference between sequence ions amino acid residues could be assigned.*

The most widely used nomenclature system for peptide fragments has been proposed by Roepstorff and Fohlman[63], where cleavage of every backbone C_α-C, C-N and N-C_α bond is covered. If the charge is retained on the N-terminal fragment it leads to the formation of A, B and C ions; if the charge is retained on the C-terminal fragment, it leads to the formation of X, Y and Z ions. The numbers presented in subscript (see Figure 1-8) indicate the number of involved residues, and is counted from the beginning of the terminal were the charge is retained. The Roepstorff-Fohlman notation has been replaced by the Biemann system[64], which features an expanded repertoire of fragment ions, including fragments that originate from the cleavage of side chains; the lettering of fragments is now in lower case letters.[65] Figure 1-8 presents the Biemann nomenclature system for all fragmentation possibilities at the peptide backbone. Note that since rearrangements such as cyclization reactions often occur, the structures of the fragment ions cannot be directly derived from the cleavage map in Figure 1-8.

The different ion activation methods mentioned at the beginning of this section lead to cleavage of the backbone at different locations and hence to the formation of different types of sequence ions. In addition, the amino acid composition, the size of the peptide, the charge state and the fragmentation time influence the fragmentation pattern.[66] Section 1.5 focusses on Collison Induced Dissociation, which mainly leads to the formation of N-terminal b-ions and C-terminal y-ions. Section 1.6 elaborates on Electron Transfer Dissociation, which leads to the formation of N-terminal c-ions and C-terminal z-ions.

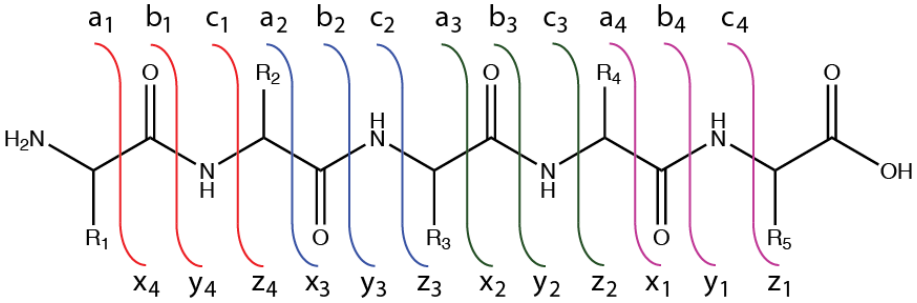


Figure 1-8. *The peptide fragmentation scheme according to the Biemann system. Counting starts from the N-terminus for a-, b- and c-fragments. For C-terminal fragments, x-, y- and z-ions, counting starts from the C-terminus.*

1.4.1. Mobile proton theory

Mass spectrometry detects only charged molecules, and in ESI-MS charging most often occurs through protonation or deprotonation. In case of adding a proton, this proton can potentially attach to the molecule at various protonation sites, such as the N-terminal amino group, the amide bond oxygen or nitrogen atom and at basic side chain groups. Usually protonation occurs at the site of highest proton affinity, but upon raising the internal energy of the molecule during collisional activation in MS/MS experiments, the proton is hypothesized to become mobilized and to be able to migrate to positions of lower proton affinity. Since dissociation is often charge-driven, the position to which the proton migrates is where cleavage occurs. This mechanism has become known as the mobile proton theory.[66-74] Mechanistic aspects determine which of the two fragments retains the added proton and thus determine the appearance of particular sequence ions in the MS/MS spectra.[66,75-77] When the added proton stays at or returns to the N-terminal residue of the peptide, N-terminal fragments (a,b,c-ions) can be expected in the resulting mass spectrum. C-terminal fragments (x,y,z-ions) are expected when the mobile proton

moves to the C-terminal residue of the peptide, for instance due to a basic residue close to the C-terminus. Neutral losses occur when the mobile proton moves to one of the amino acid residues, leading to loss of for instance H_2O and NH_3.

1.5. Collision induced dissociation

CID is the most frequently used dissociation method in bottom-up proteomics since it is relatively easily implemented on various MS platforms. Its characteristics are relatively well understood experimentally as well as mechanistically.[54,78]

1.5.1. Mechanism of collision induced dissociation

In CID, ion activation is accomplished by increasing the kinetic energy of the ion and colliding it with a neutral background gas (argon, helium or nitrogen). Translational energy is thus converted to internal energy exciting rotational and vibrational modes of the ion[79,80], which leads to statistical decomposition of the precursor ion. Consequently the cleaved bonds in CID are mostly the weakest bonds and rearrangements of the molecule prior to or during dissociation are common.[30,54,81-84]

The *pathways in competition model*[66] analyses the dissociation chemistry of peptides, by classifying the fragmentation in mechanistic, energetic or kinetic pathways. Protonated peptides dissociate by charge-remote or charge-directed fragmentation pathways. In a charge-remote reaction pathway, the charge is not directly involved in the fragmentation and the bond cleavage site is remote from the charge carrier site.[66,70] Proton transfers induce charge-directed fragmentation pathways. In the low fragmentation-energy regime, protonated peptides dissociate mainly on charge-directed pathways.[66,81] Peptides that contain aspartic acid or oxidized methionine have been shown to dissociate via a charge-remote fragmentation pathway in the low energy regime.[66] Charge-directed pathways can be divided into those that yield sequence ions (a,b,c and/or x,y,z ions) and those that yield non-sequence ions.

CID induced cleavage mainly occurs at one of the backbone amide bonds of the peptide, resulting in N-terminal b-ions and C-terminal y-ions. In addition, non-sequence ions may be formed through a number of processes; these ions likely contain valuable information the protein structure, but are hardly used in sequencing strategies.[85,86] Weak bonds in various residues may undergo cleavage and lead to neutral losses of ammonia, water and carbon dioxide. Internal fragments, having lost both N- and C-terminal ends are another common source of

non-sequence ions.[66] Furthermore, PTMs can also be detached during the CID process. Information on their position along the backbone is often lost in CID MS/MS.[32,87-91] Furthermore, sequential cleavages take place during CID, for example loss of neutral molecules from the sequence ions or the formation of smaller b- and y-ions.[92,93]

b-Type ions

Originally, b-ions were believed to possess a linear acylium ion structure produced through simple amide bond cleavage without any rearrangement.[63,64,66,94]. Cyclization reactions were considered to yield b-ion products with significantly improved stabilities than the acylium ion structure.[94] The first step in the formation of a b-ion is the migration of the added proton to one of the amide nitrogen atoms in the peptide backbone. This has the effect of weakening the amide bond and the amide carbon atom becoming electron-deficient, so that it can serve as a target of a nucleophilic attack by one of the nearby electron-rich groups.[66]

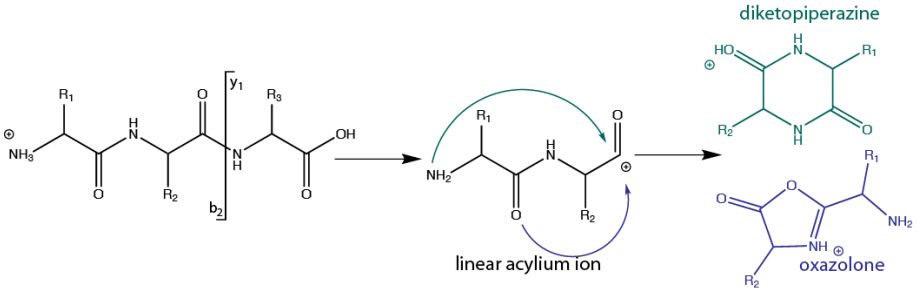


Figure 1-9. *The formation of an oxazolone or diketopiperazine structure for b_2-ions.*

The oxygen atom of the adjacent amide carbonyl group can act as a nucleophile leading to an oxazolone structure (blue in Figure 1-9).[92,95-98] Alternatively, the N-terminal nitrogen atom can act as a nucleophile leading to a diketopiperazine structure for a dipeptide (green in Figure 1-9)[99-101] or a larger macrocycle in longer b-type ions. In specific cases, hetero-atoms in a residue's side chain have been suggested to act as the nucleophile, leading to different 5- or 6-membered ring structures.[102-105] The diketopiperazine structures are lower in energy relative to the oxazolone structures. However, the trans-to-cis isomerization of the initially trans N-terminal amide bond to the cis amide bonds on the diketopiperazine structure that is necessary to form the diketopiperazine structure, imposes a significant kinetic barrier. For this reason, oxazolone structures are often kinetically favored.

For the amino acids asparagine and glutamine, the amide moieties in the side chains can be involved in the nucleophilic attack on the backbone.[66,102-105] In part II of this thesis the influence of the peptide side chain on the fragmentation mechanism is investigated for dipeptides containing Asn and Gln at the first or second residue.

a-Type ions

The elimination of a CO-molecule from a b-ion leads to the formation of a-type ions. This pathway has specifically been confirmed for b-ions with an oxazolone structure; protonation on the nitrogen atom in the oxazolone ring leads to a weakening of the -O-CO bond and eventually to elimination of the CO molecule.[66,106,107] The structure of a-type ions is observed to be an open immonium ion structure, as presented in the left panel of Figure 1-10, but cyclization reactions have also been reported.[66,106,107] The formation of a-ions has not been observed for b-ions that possess a diketopiperazine structure, and their fragmentation occurs often through the loss of an ammonia-group. Smaller b-ions can result from fragmentation of a-type ions in case of excess internal energy.[108]

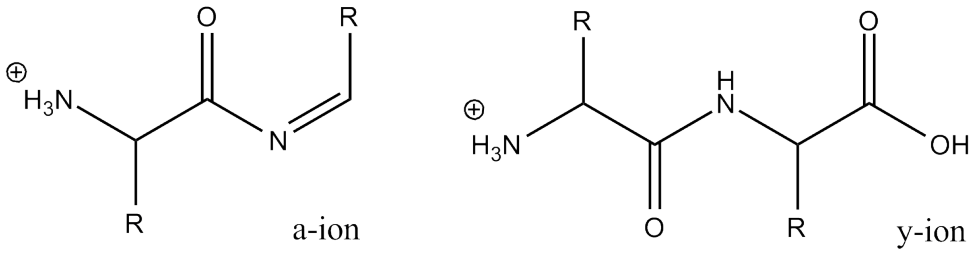


Figure 1-10. *The structures of the linear a-ion and y-ion that arise after CID.*

y-Type ions

The C-terminal counterparts of b-ions are the y-type ions. Ions of the y-type contain the added proton as well as a hydrogen that was originally attached to the N-terminal nitrogen atom.[66,95,109] The structure is that of a truncated peptide as presented in the right panel of Figure 1-10.

Neutral losses

Non-sequence ions can be formed as a consequence of expulsion of small neutral molecules such as ammonia and water [110,111] from the protonated peptide or from its sequence ions. Water can be lost from the C-terminus of the peptide, from the side chains of Asp, Glu, Ser and Thr residues or from backbone amide oxygens.[66,112-115] Loss of ammonia occurs from the side chains of Arg, Lys, Asn and Gln or from the N-terminus.[66,104,110,116,117] Loss of small neutral molecules can happen via a rearrangement / cyclization reaction, as in the formation of b-ions, or via simple bond cleavage, as is for instance often the case for species containing an amine moiety.

There are usually multiple potential pathways for the cyclization reaction to occur. As an example, Figure 1-11 presents all potential pathways and resulting isomers for NH_3 loss of a single protonated asparagine amino acid. All product ion structures are isomeric and hence have the same mass. Based on mass spectrometry alone, determining which of the reaction products is actually formed is not possible. Neutral losses from peptides containing asparagine and glutamine are investigated in part II of this thesis.

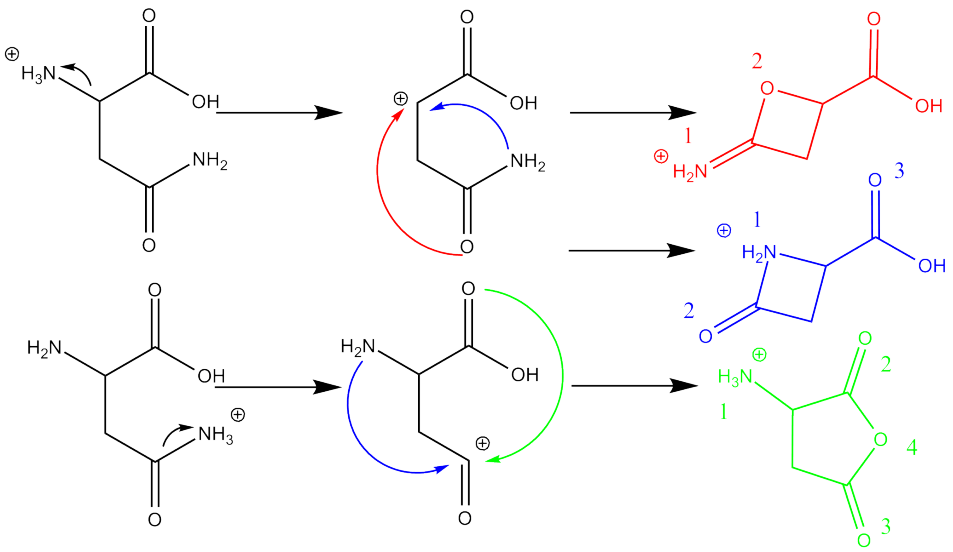


Figure 1-11. *Possible reactions for the loss of ammonia from the amino acid asparagine. The loss of ammonia also lead to several cyclization reactions and new non-linear structures. The numbers indicate the possible protonation sites.*

1.6. Electron transfer dissociation

Electron Capture Dissociation (ECD) and Electron Transfer Dissociation (ETD), jointly referred to as ExD, require a multiply charged precursor peptide, which undergoes gas-phase one-electron reduction either by electron transfer from an anionic reagent (ETD) or by capture of a free electron emitted from a cathode (ECD).[31,32,79] For both techniques, the recombination energy can induce dissociation of the peptide, which occurs predominantly at one of the N-C_α bonds in the molecule. ExD thus typically leads to the formation of c- and z-type fragment ions, where the c-ions are usually even electron and the z-ions are radicals. Besides cleavage at the N-C_α bond, losses of neutral side chain molecules appear and charge reduced intact peptide ions are observed, which result from non-dissociative electron transfer (ETnoD).[118] Obviously, to generate charged fragments observable in ETD, the precursor ion needs to be multiply charged. Anthracene[59,119], 9,10-diphenyl-anthracene[119], azobenzene[120,121], azulene[120] and fluoranthene[121-123] are suitable as anionic reagents in ETD.[120,124,125]

1.6.1. Methods of electron transfer dissociation

The first ECD experiment was reported in 1998 by McLafferty and coworkers[58], who modified the ion cyclotron resonance (ICR) cell to include a dispenser cathode to irradiate the stored ions with low-energy electrons. Charge-reduced species and c- and z-ions were observed, indicating that peptide ion fragmentation pathways were fundamentally different from those known from CID reactions.[58,79] But, this was not the first time that c- and z-type ions were observed. In 1996 Guan et al. applied ultraviolet irradiation on multiply-charged cations leading to these unusual fragments.[126] The formation of these ions was caused by the secondary electrons that were generated by the UV laser light hitting the metal ICR cell rather than the UV photons themselves.[58,79,126]

Not all types of MS instruments enable successful and stable application of ECD.[79] Due to low mass cut-off restrictions in RF ion traps, electrons cannot be stored along with molecular ions that are several orders of magnitude higher in mass. To overcome this restriction, Syka, Coon and Hunt[59] used anions as electron donors to transfer electrons to multiply-charged peptide cations, creating the electron transfer dissociation technique. At that time, McLuckey and coworkers had already successfully shown simultaneous trapping of positively and negatively charged ions in a quadrupole ion trap to drive and investigate ion/ion reactions.[127]

The formation of c- and z-ions instead of b- and y-ions in ExD experiments is believed to be due to two important differences.[128] First, the peptide bond fragmentation in CID occurs through intermediates that require relatively long times (up to microseconds) to form.[129] These times are long compared to the rate of intra-molecular vibrational energy redistribution (IVR) (for a detailed explanation see Section 2.1), so that peptide-bond fragmentation is usually ergodic. In contrast, ExD is believed to occur much faster ($<10^{-12}$ s), which allows only for direct bond cleavage to happen in a non-ergodic process[128], although this non-ergodicity has been disputed.[130,131] Moreover, the ExD fragmenting species are not even-electron ions but hydrogen-abundant radical cations.[128] The presence of a radical site changes the strengths of the adjacent bonds[128], and N-C_α bond cleavage can be explained by radical-induced dissociation following the capture of the excess hydrogen atom by the carbonyl oxygen. Alternative mechanisms for N-C_α bond cleavage have also been proposed and section 1.6.2. presents a summary of the mechanisms that have received most attention in the literature.

ECD versus ETD

Since the observed fragmentation patterns for ECD and ETD are similar, their dissociation mechanisms are often assumed to be identical. However, there are some differences between the two techniques.[59] The amount of energy added to the molecule is larger for ECD than for ETD. This results in the formation of additional peaks in ECD mass spectra, for instance due to additional neutral loss channels. In addition, ECD gives more charge-reduced (ETnoD) species than ETD. ECD experiments are performed under the high vacuum (~10^{-10} Torr) conditions of a FTICR instrument, while ETD reactions take place in ion traps where the pressure is several orders of magnitude higher.[31] The higher pressures result in faster collisional cooling and a reduction of the internal energy of the ion that attached the electron and reducing the occurrence of more energetic fragmentation routes. Furthermore, the pressure difference influences the relative abundance of the charge-reduced species, the number of peaks attributable to neutral losses and the radical/even-electron ratios of fragment ions.[31]

The application of ETD in proteomics

ExD enables top-down mass spectrometry of intact proteins since it yields more extended protein fragmentation.[31,124,132,133] ExD dissociation particularly results in a larger number of cleavages in the middle portion of the protein, whereas CID often yields cleavages mainly near the C- and N-terminal ends. ExD thus leads to the formation of more sequence fragments and hence better sequence coverage than

CID, improving database matching scores. However, the top-down approach requires high-resolution mass spectrometers. On the other hand, ETD is less suited for bottom-up proteomics, because of the low charge states of the small tryptic peptides, making dissociation induced by recombination less efficient. ExD is better suited for a middle-down approach, in which the peptides are longer and have a higher charge state.[31]

A major advantage of ExD over CID in the analysis of PTMs is the fact that ETD keeps the labile modifications intact at the modified residues. Because in ExD intact proteins can be sequenced, loss of PTMs during the enzymatic digestion of peptides is avoided.[124] ETD MS/MS can localize the exact site of the PTM, which may be lost using CID MS/MS.[78,87,91,128,134,135] Furthermore, ExD proved to be helpful in distinguishing the isomeric amino acid residues leucine (Leu) and isoleucine (Ile) through the formation of different secondary fragments from each of the two residues.[136-140]

CID and ExD are complementary in protein identification as CID generates b- and y-type ions while ETD produces c- and z-type ions. Both methods suffer from an unpredictability for fragmentation events, which makes *de novo* sequencing challenging.[78,124,138] A combination of CID and ETD in protein identification has been shown to outperform one of these techniques alone.[78,124,138]

1.6.2. ExD peptide fragmentation mechanism

The precise reaction mechanisms of ETD dissociation have been the subject of extensive discussion.[31,58,59,79,125,128,130,141-156] The first mechanism, proposed by McLafferty et al.[58], considered electron attachment to a protonated backbone amide group to form an amino ketyl radical intermediate that undergoes a β-fission involving the N-Cα bond on the C-terminal side of the affected amide group. This mechanism was considered unlikely because the backbone amide groups are less basic than the side-chains of Lys, Arg and His residues. In a mechanism proposed later by McLafferty (Cornell mechanism), the electron is captured by a charged site in the molecule, commonly an ammonium NH_3^+-group, which is hydrogen bonded to a nearby backbone amide carbonyl.[141,146,156] Hydrogen transfer from the protonation site to the carbonyl group then induces backbone N-C_α bond cleavage and leads to c-type ions that possess an enol-imine structure.

Although the Cornell mechanism adequately explains the observed fragments, not all observed phenomena during the ExD process could be explained.[31,125] This led to

variations on the Cornell mechanism.[152,156,157] An example is the Uppsala mechanism proposed by Zubarev et al.[152] In this mechanism, electron capture occurs at a neutral hydrogen bond between the backbone NH and carbonyl groups, resulting in a negative charge on the nitrogen atom and an aminoketyl radical after a hydrogen transfer. The cleavage process that follows is the same as in the Cornell mechanism. The mechanisms named before are "right-side" dissociation pathways: cleavage of the N-C_α bond occurs to the right side of the carbonyl group. An alternative pathway derived from the Cornell mechanism is the "enol-mechanism"[154], where the N-C_α bond is cleaved at the left side of the carbonyl group.[153] Figure 1-12 shows a schematic overview of the proposed fragmentation mechanisms for the ETD-reaction in peptides.

A new model was introduced to overcome shortcomings in (variations of) the Cornell-mechanism[125], such as the relatively high energy and possible instability of the resulting enol-imine form of the c-type fragments. Furthermore, theoretical modelling of the electron capture process suggested that the ammonium groups may not be the most likely sites for initial electron capture. The Utah-Washington mechanism[143,144,149,158] is now the most widely accepted ExD fragmentation mechanism. This mechanism was originally proposed by Simons et al.[144] at the University of Utah and further extended by Tureček et al. from the University of Washington.[143,159] The electron is initially captured in an empty π^* orbital of a backbone amide carbonyl. A charge-stabilized amide anion-radical intermediate results, which isomerizes by proton migration finally leading to cleavage at the N-C_α bond. The c-type ions in this mechanism are formed directly with an amide group at its C-terminal end (see Figure 1-12). Within this mechanism, discussions are going on about how the electron is transferred to or captured by the amide π^* orbital, and the order of steps (first proton transfer or first bond cleavage) is under debate.[31] For this reason refinements of this most accepted mechanism are examined in literature.[160]

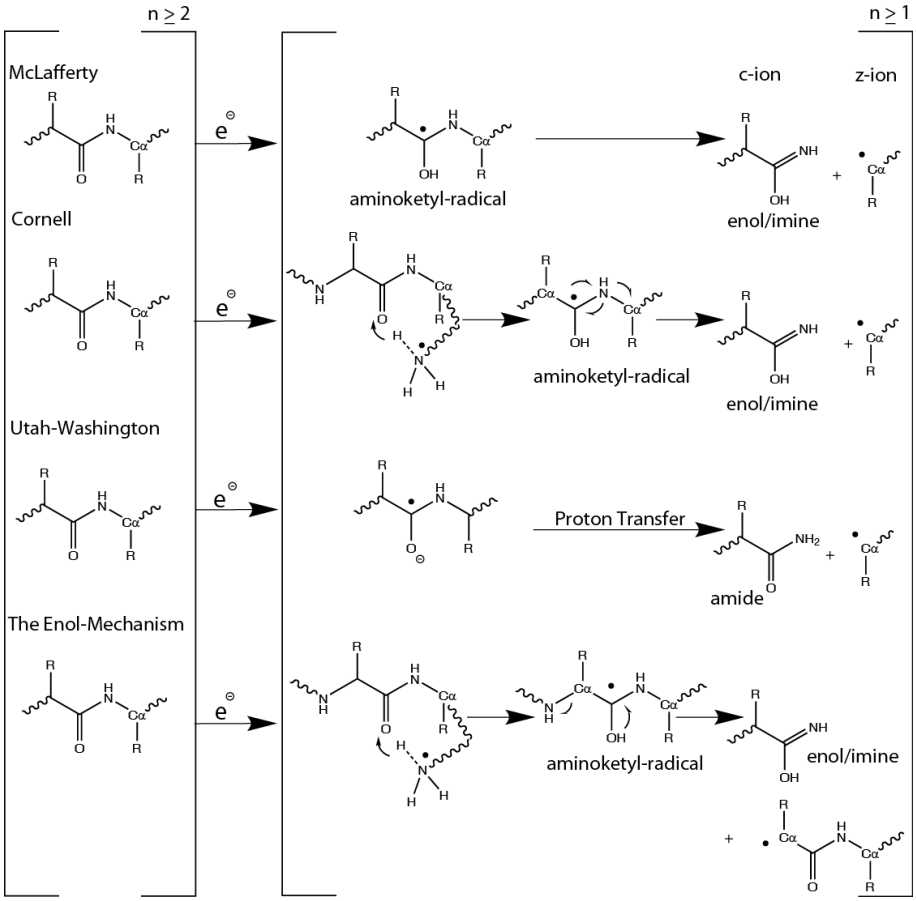


Figure 1-12. *Overview of the most popular proposed mechanisms for ETD reactions. The first mechanism was proposed by McLafferty and has been adapted over time.*

c-Type ions

Figure 1-12 presents two pathways for the formation of c-ions after N-C_α-bond cleavage, which leads to an enol-imine structure (Cornell mechanism) or amide structure (Utah-Washington mechanism). Possibly, the enol-imine structure is formed first, but subsequently rearranges to the thermodynamically more stable amide. In one possible pathway an anionic –C(O^-)=NH group is formed first at the C-terminus of the c-ion, which then abstracts a proton to form either an enol-imine or an amide. Alternatively, the proton is abstracted prior to cleavage of the N-C_α bond, in which case formation of the enol-imine would be favored. A more detailed description of the formation and structures of c-ions is given in Chapter 8. This

chapter also presents the structures of a set of 10 c-type ions that have been elucidated by infrared ion spectroscopy.

z-Type ions

During the ETD process radical z-ions are formed complementary to the formation of c-ions.[161] In Figure 1-12 the radical is positioned at the α-carbon of the cleavage site. However the possibility is that hydrogen migration occurs and that the radical site migrates in the ion.[162] There are two options for radical migration, first the radical migrates to the β- or γ-carbon within the amino acid residue at the cleavage site.[163] The second option is radical migration to α-, β- or γ-carbon along the peptide backbone.[164-168] Using infrared ion spectroscopy, the position of the radical has been investigated for six z-ions. The results are presented in Chapter 9.

Side chain losses

Every amino acid residue has its unique side chain that can be detached during ExD activation. Except for Gly and Ala residues, losses of neutral fragments are commonly observed.[53,169-175] Some neutral losses occur immediately upon ExD, while others are typically observed as secondary dissociations that are only formed upon applying additional CID activation on a radical z-ion (see Chapter 10).

Two types of side chain losses are identified, one occurring directly at the cleavage site that formed the z-ion, leading to linear w-type ions, the other one remote from the primary backbone cleavage site leading to cyclic u-type ions.[176] Figure 1-13 shows the suggested reaction mechanisms involved in the formation of u- and w-type ions.[128,137,175] Before w-type ions are formed, the radical that is originally located at the α-carbon of the cleavage site migrates to the C_β or C_γ position of the same residue.[53,125,137,140,172,176] Direct formation of w-type ions then proceeds by cleavage of the C_β-C_γ bond and the formation of a double bond between the C_α and C_β atoms. A neutral radical fragment is expelled and an even-electron w-type ion results. For the formation of u-type ions, the radical migrates from its original site to the C_β or C_γ atom in the adjacent residue.[53,125,137,140,172,176] Bond formation between the N-terminal C_α atom and the C_β atom of the adjacent residue produces a cyclic u-type ion.[137,175-177]

In addition to losses of parts of the side chain as described above, loss of the entire side chain of an amino acid residue has also been observed during ExD.[175,176] In this case, the radical remains on the C_α-position of the peptide. A possible mechanism for this reaction involves abstraction of a hydrogen atom from the C_γ-position in the residue. In the subsequent step the C_α-C_β bond is cleaved and the side chain is expelled as an even-electron neutral species.[175,176] This mechanism is depicted in the bottom panel of Figure 1-13.

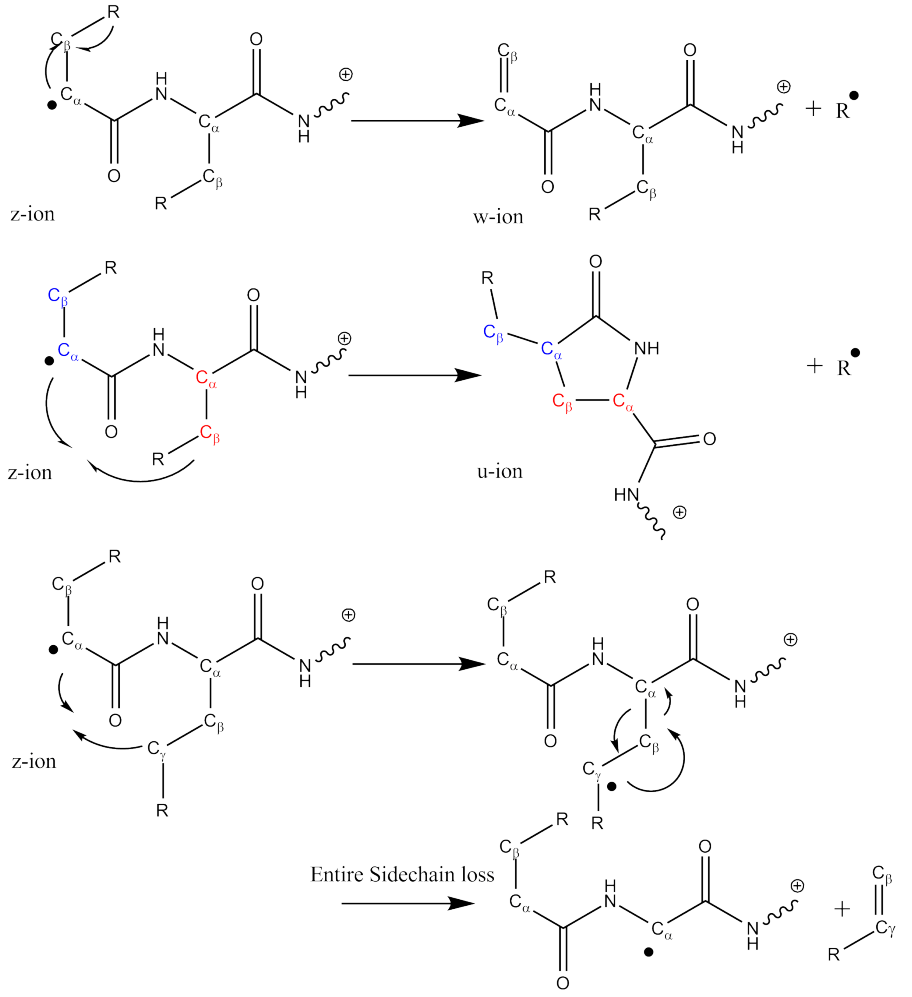


Figure 1-13. *The reaction mechanism for the formation of u- and w-ion type formation by loss of a neutral fragment and the reaction mechanism for the formation of the entire side chain loss (H-atoms are implicit).*

1.7. Fragment structure determination

Detailed information about the amino acid sequence in proteins can be obtained by tandem mass spectrometry and analyzing the m/z values of the MS/MS fragment ions. The molecular structures of the fragment ions, however, often remain poorly known. Based on the m/z value alone, discrimination between possible product structures is usually not possible. Even if the chemical formula of the ion of interest is known, many isomers can usually be suggested to correspond to this chemical composition. Since mass spectrometry involves by definition charged species, the various possible (de)protonation sites further complicates a precise structure determination. Furthermore, a peptide molecule is "floppy" and can adopt different three-dimensional configurations (conformers) by rotation around the dihedral angles. Several methods have been developed to resolve the three-dimensional structure of ions in mass spectrometry, including ion-mobility [178], H/D exchange [179], isotopic labeling [180] and infrared ion spectroscopy[181].

The IR spectrum contains valuable information to derive structural properties of the ionized molecule. Chemical moieties such as carbonyl, carboxyl and amide groups possess diagnostic vibrational modes, which can be probed by infrared ion spectroscopy (for further reading see Chapter 2.1). Ion spectroscopy has successfully been applied to determine the structures of precursor ions including protonated peptides[182] and amino acids[71,183,184], but also of MS/MS (peptide) fragments ions.[161,185-193] Furthermore, the structures of radical ions have been addressed and elucidated.[194-197] Particularly for small-molecule mass spectrometry, infrared ion spectroscopy thus is a valuable method to distinguish between isomers, protomers and conformers[74,107,198-206]; applications to larger biopolymers, such as especially proteins are scarcer but in development[207,208]. Structural information is often derived from an IR spectrum with the use of theoretical predictions based on high-level quantum-chemical calculations (see Chapter 2.3), but vice-versa, the experimental spectra can also be employed to benchmark the performance of theoretical methods in prediction IR spectral properties for ionized species.

As an example of the use of IR ion spectroscopy to determine the structure of peptide MS/MS fragments, we mention here the assignment of an oxazolone structure for the b_2 ion originating from the protonated tripeptide $[GlyGlyGly+H]^+$. The oxazolone structure is unambigously established from the diagnostic oxazolone ring C=O stretch near 1960 cm^{-1} in the IR spectrum of this mass-isolated MS/MS

fragment [209], and on the basis of the close agreement between experimental and predicted spectra in general. Oxazolone b_2 ions were also spectroscopically established for $[AlaGlyGly+H]^+$ [191], $[AlaAlaAla+H]^+$ [190] and others, firmly cementing the importance of this chemical structure in peptide CID MS/MS. But exceptions were also encountered: the b_2 ions of $[PheGlnAla+H]^+$ and $[AlaAsnAla]^+$ were spectroscopically shown to possess glutarimide and succinimide structures, respectively, as a consequence of side-chain involvement in the reaction mechanism.[102] A small number of studies so far have addressed the spectroscopic determination of ExD MS/MS fragment structures including a c-ion[189], radical z-ions[161] and w-ions[210]. More applications of IR ion spectroscopy are highlighted in several review articles.[187,200,211,212]

1.8. Outline of this thesis

In this thesis, a new MS platform that was recently coupled to the FELIX IR laser, a modified Bruker Amazon Speed ETD Ion Trap mass spectrometer, has been utilized. This new instrument allowed us to measure Infrared Multiple Photon Dissociation (IRMPD) spectra more efficiently and with a higher sensitivity, enabling us for instance to measure a large number of IR spectra for MS, MS^2 and MS^3-species. Resolving their structures generates an accurate structural map of a peptide fragmentation tree, as we explore in Part II of this thesis. In particular, CID has been applied on the protonated amino acids asparagine and glutamine and on four protonated dipeptides containing Asn and Gln residues. We spectroscopically investigate deamidation and dehydration reaction products and propose the mechanisms for these reactions. Despite the close similarity of the precursor ions, a large variety in the reaction mechanisms is observed.

The Bruker Amazon is equipped with a radical anion source to enable ETD MS/MS experiments. In this sense, the coupling of an ETD enabled mass spectrometer with an IR free electron laser is unique. It enables for the first time an extensive spectroscopic investigation of ETD MS/MS fragment ions to elucidate their structures and to address ETD reaction mechanisms, which have been under much debate over the past decades. In Part III of this thesis, spectroscopic investigations of extensive sets of c-, z-, and w-type ETD fragment ions are presented. From the large number of spectra recorded, a picture emerges in which c-ions appear to possess mainly (or exclusively) amide c-type structures and in which the radical in z-type ions appears not to migrate in most cases.

Chapter 2

Experimental and theoretical methods

In this thesis, IR spectroscopy is used to derive molecular structure information. Absorption bands in the IR spectrum of a molecule relate to the normal modes of vibration of the molecule, *i.e.* to the oscillatory motion of the atoms constituting the molecule with respect to each other. The precise values of the vibrational frequencies depend sensitively on the bond strengths and atom types in a molecule. Molecular structure and IR absorption frequencies are thus intimately related and retrieving the molecular structure from the observed IR spectrum lies at the heart of the studies presented herein.

2.1. Infrared spectroscopy

Classically, IR absorption bands at specific frequencies are correlated with specific functional groups in a molecule. Table 2.1 lists characteristic IR frequencies for a series of typical vibrations in specific functional groups that occur in organic molecules in general, and in the peptides under study in this thesis in particular. The table also includes an indication of the strength of the absorption. More advanced correlations between experimental IR spectra and accurate molecular structures can be obtained from quantum-chemically predicted IR spectra as will be discussed in more detail below.

Recording the IR spectra of molecules is classically done by a spectrophotometer, in which the attenuation of a light beam transmitted through the sample is measured as function of the wavelength. According to the Lambert-Beer law, the absorbed light by a solution is related to the thickness of the absorbing layer and the concentration of the absorbing molecule according to equation [2-1]:

$$log\frac{I_0}{I} = \varepsilon_\lambda c l \qquad [2\text{-}1]$$

I_0 is the intensity of the incident light, I is the intensity of the transmitted light, the ratio I/I_0 is the transmittance, ε is the molar extinction coefficient which is λ-dependent, c is the concentration of the absorbing species and l is the path length of the light-absorbing sample.[2,213] The molar extinction coefficient is an intrinsic molecular property and the absorption length is usually determined by the spectrometer design. Therefore, in practice one adjusts the concentration – or the pressure in the case of a gaseous sample – to obtain an observable light attenuation. However, choosing the required number density is only possible for neutral gaseous molecules. For ionized species, such as the ions inside a mass spectrometer that are of interest here, the density is extremely low due to mutual Coulombic repulsion. The use of a classical spectrophotometer to record the IR spectra of ions inside a mass spectrometer is therefore impossible. To resolve this problem, a form of *action spectroscopy* is applied to access the IR spectra of mass-isolated ions in the mass spectrometer. The photon absorbed by a molecular ion induce photodissociation – the action – of the ion, which is monitored in the mass spectrum.

Characteristic Vibration		Wavenumber (cm^{-1})	Peak Intensity
O-H Stretch	Alcohol, phenol	3525 – 3200	Strong
	Carboxylic Acid	3400 – 2500	Strong
N-H Stretch	Primary Amine	3400 – 3500	Medium/Weak
	Secondary Amine	3500 – 3300	Weak
	Amide	3350, 3175 - 3150	Weak
C-H Stretch	Alkene	3100 – 3000	Medium/Strong
	Aromatic	3100 – 3000	Medium/Strong
	Alkane	3000 – 2800	Medium/Strong
C=O Stretch	Aldehyde	1740 – 1690	Strong
	Ketone	1750 – 1780	Strong
	Carboxylic Acid	1780 – 1700	Strong
	Amide	1690 – 1630	Strong
C=C Stretch	Alkene	1670 – 1600	Weak/Medium
	Aromatic	1600 – 1575 1500 - 1450	Medium
N-H Bending	Amide	1655 – 1610	Medium
	Primary Amine	1650 – 1560	Medium/Weak
C-H Bending	Alkane	1470 – 1370	Weak/Medium
C-N Stretch	Amide	1410	Medium
	Aliphatic amine	1250 - 1010	Weak/Medium
	Aromatic amine	1370 - 1250	Weak/Medium
C-O Stretch	Carboxylic Acid	1320 – 1210	Medium
	Alcohol, Phenol	1255 – 1000	Weak
C-H Bending (out of plane)	Aromatic	900 – 675	Weak
N-H Bending (out of plane)	Amine	910 – 660	
	Amide	800 – 670	Weak

Table 2-1. *List of characteristic IR absorptions in peptide fragment ions.*[214]

For many spectroscopic applications where the density of molecules is low (*c* in Equation 1), different methods of action spectroscopy have been developed. Infrared Multiple-Photon Dissociation (IRMPD) spectroscopy[211,215-219] is the method of choice when one is interested in recording IR spectra of ions in a mass spectrometer. This technique is based on the sequential absorption of multiple photons by a single molecular ion so that dissociation is induced. The technique is further explained below. Other examples of IR action spectroscopy methods are Ion Dip Spectroscopy[213,220,221] and Messenger Spectroscopy[222-224].

2.1.1. IRMPD spectroscopy

IRMPD spectroscopy has proven to be a valuable method to obtain IR spectra of isolated ions in a mass spectrometer. The IRMPD process initiates with the absorption of a resonant photon by the molecule. This photon is resonant with the excitation from the vibrational ground state to the first excited state of a specific normal mode in the molecule. To reach the dissociation threshold, additional IR photons need to be absorbed. A typical bond that is cleaved during the IRMPD process is for instance a carbon-nitrogen bond, which requires about 305 kJ/mol of energy.[225] The energy of a 10 µm photon is 12 kJ/mol, so that at least 25 photons are needed to cleave the C-N bond. However, in practice more photons are required to break this bond because of (collisional) deactivation processes that occur at the same time. Collisional deactivation by the helium buffer gas present in the ion trap can be reduced to some extent by using short-pulse lasers so that the excitation rate outcompetes the deactivation processes and prevents the molecule from cooling down.[181] Reducing the background pressure in the trap also increases the excitation efficiency, although it may reduce the quality and the intensity of the mass peaks in the MS.[226]

Figure 2-1 presents the typical vibrational potential of a normal mode within a molecule, which shows that the distance between the energy levels becomes smaller as the energy increases. This anharmonic shape of the vibrational potential prevents sequential photon absorption in a ladder-climbing fashion. After the absorption of a resonant photon, the laser frequency is no longer resonant with subsequent transitions in the potential. This effect is referred to as the anharmonicity bottleneck for monochromatic laser sources. Because of this bottleneck, the dissociation limit cannot be reached within this potential well. At the same time, anharmonicity also introduces couplings between different vibrational modes in the molecule, which facilitates a fast energy transfer from the absorbing mode to the bath of all other vibrational degrees of freedom. This

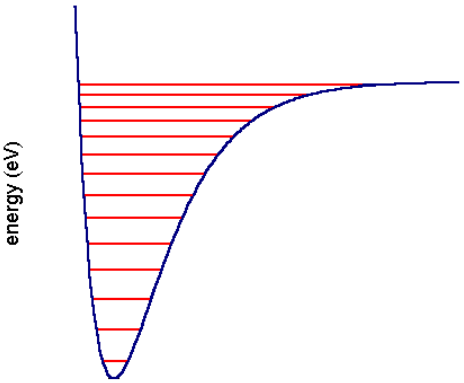


Figure 2-1. *The vibrational potential of a molecule is anharmonic, the distance between the levels becomes smaller at higher energies.*

dissipation of energy into the bath is known as intramolecular vibrational redistribution (IVR). When an excited vibrational level is deactivated by IVR to other modes, another photon can be absorbed. This process can be repeated multiple times, so that the molecule is progressively energized until the internal energy of the molecule is raised to above the dissociation limit.[187]

For the experiments described in this thesis, the dissociation of the molecule is measured as a function of the IR frequency. Mass spectrometry is able to detect the change in mass to charge ratio of the ion (this includes the change in charge of multiple charged ions) due to the dissociation process. When the laser is off resonance, no IR photons are absorbed and no fragmentation is observed (see top panel Figure 2-2). The IRMPD fragmentation yield is zero at this frequency. When the laser is tuned to a frequency that is on-resonance with an allowed vibrational transition, the ion absorbs multiple infrared photons and undergoes dissociation (see bottom panel of Figure 2-2). The IRMPD yield is calculated using equation [2-2]: the sum of the intensities of the observed fragments is divided by the sum of the intensities of all ions, *i.e.* precursor and fragment ions.

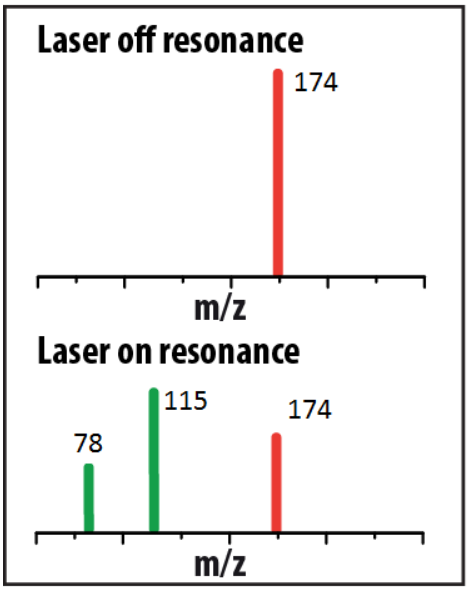


Figure 2-2. *When the isolated molecule of interest (m/z 174) absorbs the IR photons, the dissociation limit is reached and smaller fragments will appear in the mass spectrum (m/z 78 and m/z 115). To obtain an IRMPD spectrum the fragment yield is plotted against the wavenumber.*

$$\text{IRMPD yield: } \frac{\sum Intensity\ of\ fragments}{\sum Intensity\ of\ fragments + Intensity\ of\ precursor\ ion} \qquad [2\text{-}2]$$

A disadvantage of using IRMPD spectroscopy is that spectral bands can be broadened, red-shifted or affected in relative intensity[181,227-232] as a consequence of the multiple-photon excitation mechanism. These effects are generally related to the anharmonic nature of the vibrational modes of any molecule. As the internal

energy of the molecule increases during the multiple-photon excitation, band frequencies shift leading to an overall band broadening. Modes with very large anharmonicities may particularly be affected; resonance with the laser frequency may be lost during excitation, which then affects the observed intensity of this band. However, computed linear absorption spectra typically provide a good approximation of IRMPD spectra, so that they can be used to determine the molecular structure of the ion of interest. Caution is especially required for modes with shallow or double-well potentials, such as shared-proton vibrations or NH bending modes of amines and amides.

2.2. Experimental implementation

To be able to dissociate a covalently-bound molecular ion and to measure its IRMPD spectrum a powerful and widely tunable IR laser is required. In ion traps with a relatively high background pressure, such as quadrupole ion traps, a pulsed laser is preferred over cw laser sources because the intensities during the pulse lead to excitation rates that can outcompete collisional deactivation. A wide wavelength tunability is required to cover a sufficiently large fraction of the spectrum, which is necessary for a reliable structure determination. The Free Electron Laser for Infrared Experiments (FELIX) used in the experiments described in this thesis possesses exactly these properties. To measure the infrared spectrum of the isolated (fragment) ion of interest, an optically accessible ion trap mass spectrometer is coupled to the beam line of FELIX.

2.2.1. The ion trap mass spectrometer

For all MS experiments described in this thesis a Bruker AmaZon speed ETD 3D radio-frequency ion trap mass spectrometer has been used.[233] A schematic cut through the instrument is shown in Figure 2-3. Charged peptides are created in the ion source and are guided to the ion trap module, where positively or negatively charged ions can be stored, mass-isolated and the mass-to-charge ratio can be determined.[234-237] Two holes are made in the ring electrode of the trap to allow the FELIX laser beam to access the stored ions and finally to measure their IRMPD spectrum.[226]

Ions are generated in the ion source (see section Electrospray Ionization) and then led into the quadrupole ion trap via a series of ion guides and funnels[233] that guide charged species but remove neutral gas and solvent molecules from the ion path.[233-236] The ions are gently pushed towards the entrance of the ion trap module by a DC gradient. As the ions are continuously produced, the flow of ions is

interrupted during mass analysis and IRMPD spectroscopy, using a gate lens and a focusing lens assembly.[233] Ions can freely move into the trap during the accumulation time but are extracted out of the transfer path during the period when the ions are trapped.

The Ion Trap Module consists of three main components: the pre-trap ion conditioning unit, the (near-)quadrupolar ion trap and a detector unit.[233-236,238] The quadrupolar ion trap consists of three electrodes, of which the two endcap electrodes have an identical hyperbolic shape. Both end-cap electrodes have small perforations in the center for the ions to enter the trap and to exit the trap to the detector. The third electrode is a donut-shaped ring electrode with hyperbolic inner surface, which is positioned in between the two end-cap electrodes.[234-236,238] A schematic overview of the three electrodes and a cut through the ion trap in presented in Figure 2-4. The pressure in the trap is around 10^{-3} mbar, which is maintained by a turbomolecular pump. For efficient trapping and cooling of the generated sample ions, a helium buffer gas is introduced into the ion trap.[233]

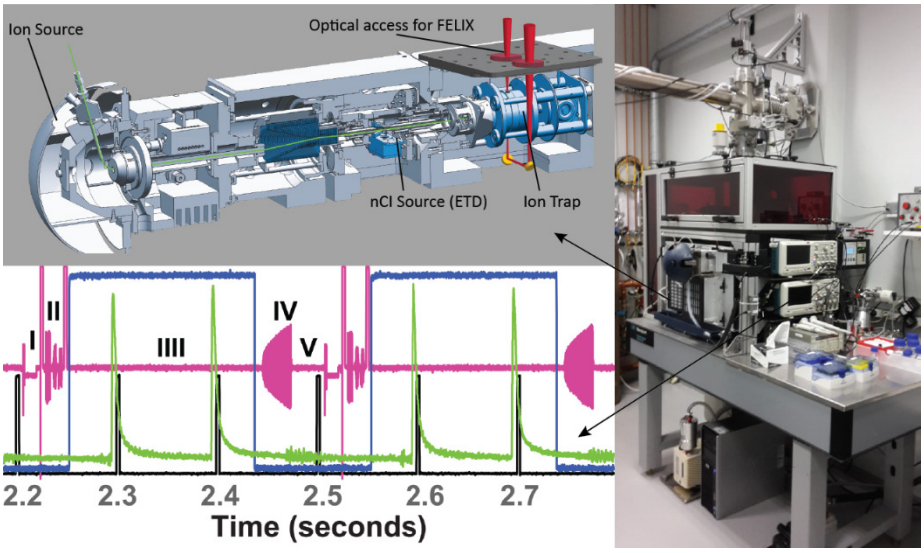


Figure 2-3. *An overview of the experimental set up in the lab (left panel), the inside of the ion trap (top right panel) and the oscilloscope window which monitors the experiment (bottom right panel). Partly adapted from*[226].

To accumulate the ions in the trap, a high voltage radio frequency (RF) potential on the ring electrode is applied leading to a potential well where ions are confined. The end cap electrodes are held at ground voltage. Depending on the amplitude of the RF voltage, the field can trap ions of a particular mass range.[233-236,238] The mass/charge ratios of the trapped ions are measured when they are scanned out of the trap and focused onto a conversion dynode detector. Ramping the RF amplitude, the ions leave the trap potential well in order of ascending mass/charge ratio and impinge on the detector creating a series of ion signals dispersed in time, which constitutes a mass spectrum.[233-236,238] The detector is equipped with an ion gate, which prevents the ejected ions during isolation or fragmentation steps from reaching the detector.[233] Ejection of ions from the trap potential well is accomplished by ramping in a linear fashion the RF amplitude applied to the ring electrode, with additional RF signals being applied to the endcap electrodes.[2,233-236,238]

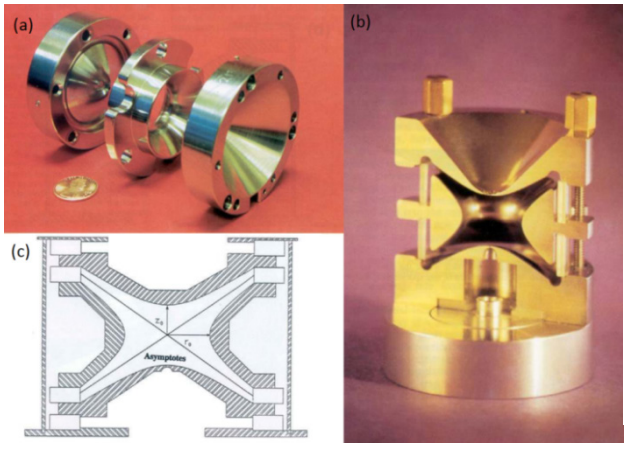


Figure 2-4. *Panel (a) shows the three electrodes of the quadrupole ion trap, panel (b) shows the ion trap cut in half along the axis of cylindrical symmetry, panel (c) shows a schematic diagram of the three-dimensional ion trap including the asymptotes. Figure adapted from [235].*

Collision induced dissociation in the ion trap

The ion trap not only stores ions, it can also be programmed to induce dissociation by CID in MS/MS or MS^n experiments. The trapped ions continuously undergo collisions as they move along their trajectories through the trap in the oscillating quadrupole field generated by the RF drive voltage. By applying tailored RF pulses to the endcap electrodes, ions of a specific m/z value can be resonantly accelerated, increasing their kinetic energy. These ions will undergo more energetic collisions with the helium buffer gas in the trap, raising their internal energy. The kinetic energy and thereby the collision energy can be varied, so that different degrees of collisional excitation can be induced.[2,234,236]

In a tandem mass spectrometry (MS/MS or MS^2) experiment, the molecule of interest is usually isolated by selecting its mass-to-charge ratio and expelling all other ions from the trap; the ion of interest is then selectively dissociation by CID. The instrument is capable of sequentially performing multiple isolation and CID steps, enabling MS^n experiments, in which products of products of products etc. can be studied. In this thesis, isolated precursor ions were subjected to CID for 40 ms using an amplitude parameter in the Bruker software of approximately 0.3-0.5 V.

Electron transfer dissociation in the ion trap

As an alternative to CID, ions can also be dissociated using ETD, as explained in Chapter 1. The Bruker Amazon instrument possesses a negative Chemical Ionization (nCI) source, which produces the reagent anions for ETD reactions. The nCI-source is located after the ion funnel for the analyte ions, form where the anions are transported to the quadrupole ion trap via the same hexapole ion guide that is used for the analyte ions (see Figure 2-3). The radical anion of fluoranthene is generated in the nCI source from sublimed solid fluoranthene supplied from a cartridge and a heated crucible in the instrument; methane enters the nCI source via an external inlet. A chemical ionization plasma is created in the nCI source via a cascade of reactions. The plasma contains low-energy electrons and electron attachment of these thermal electrons to the neutral molecules of fluoranthene leads to the production of the radical anion.[239]

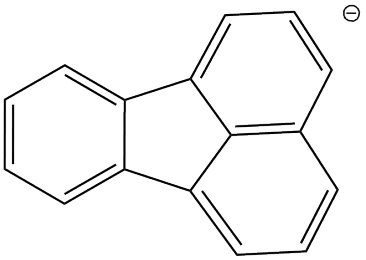


Figure 2-5. *Structure of the fluoranthene radical anion used for the ETD reaction.*

The radical fluoranthene anions (structure is shown in Figure 2-5) are guided into the ion trap and stored along with the isolated multiply charged precursor peptide ion of interest.[233,239,240] Equation [2-3] presents the chemical equation for the electron capture process.

$$[M + 2H]^{2+} + A^{-\bullet} \rightarrow [M + 2H]^{+\bullet} + A \qquad [2\text{-}3]$$

Over the course of typically 50-300 ms, an electron transfer reaction takes place reducing and dissociating the peptide ion.[240] The dissociation time is optimized to

obtain the highest intensity of the ETD-fragment of interest for IRMPD spectroscopy.

IRMPD spectroscopy in the ion trap

Optical access to the ion cloud is accomplished by modifying the ring electrode introducing 3-mm holes in its top and bottom. Infrared transparent windows (KRS-5) were positioned in the vacuum housing exactly above the ion trap and two gold-coated mirrors were placed below the trap so that the laser beam of FELIX can pass through the trap and be guided back out of the vacuum chamber (see Figure 2-3).[226] A pyroelectric energy sensor (Thorlabs, model ES220C) measures the pulse energy of the light returning from the ion trap. The optical box containing the mirrors shaping and directing the beam into the ion trap is flushed with dry nitrogen gas to avoid attenuation of the IR beam by water vapor and carbon dioxide.

The inset of Figure 2-3, shows a typical oscilloscope window monitoring the experimental sequence. The black line is the master trigger for the experiment, generated by the optical FELIX pulse, which starts the MS sequence for the measurement of a spectrum. The pink trace is picked up from the RF drive voltage of the ion trap and enables visualization of, and synchronization with, MS events, such as ion accumulation (I), ion isolation (II), scanning the ions out of the trap (IV), and waiting time until the next laser trigger to start the following sequence (V). The flat sections in the pink trace (III) indicate the times where IRMPD can take place (IRMPD window). The blue trace indicates the opening time of the optical shutter, which is placed in the IR beam path in front of the ion trap. When the blue trace is high, the shutter is open, enabling the FELIX beam to enter the trap. The green trace is energy of the laser pulse monitored by the sensor placed at the exit of the ion trap.[226]

Electrospray ionization

Studying biological macro molecules such as peptides and proteins by mass spectrometry has only become possible with the invention of Electrospray Ionization (ESI)[241] and matrix-assisted laser desorption ionization (MALDI)[29]. ESI is a nondestructive soft-ionization technique, which is able to produce intact positively and negatively charged ions from a sample solution[28,242]. Our instrument is equipped with a Bruker Apollo ESI source, shown in Figure 2-6. During ESI, charged microdroplets containing the peptides to be studied are sprayed into the spray chamber (2) through the ESI sprayer (1), which ionizes the droplets that exit the tip.

The needle is kept at a high electrical potential difference with respect to the inlet of the mass spectrometer, dispersing the solution into a fine mist of charged microdroplets (nebulization or aerosol generation). A counterflow of heated drying gas (N_2) is used to aid in the evaporation of the solvent and thus in the volatilization of the analyte ions. As the diameter of the droplets decreases due to evaporation, the surface charges are forced closer together[242]. When the concentration of charges is so high that the Coulombic forces overcome the surface tension of the droplet, dispersion of the drop into smaller droplets occurs. These droplets continue to evaporate until all solvent has evaporated and the charged analyte remains.[2] Uncharged molecules are carried away to the waste (4) container in the ESI source.[233]

For positive ions, the spray assembly (1) is held at ground potential and ions are focused on the entrance of the capillary held 3-5 kV below the sprayer and 200-500 V below the spray shield (3) directly in front of the capillary entrance.[226] Ions are drawn into the capillary when near the entrance and are further transported to the exit of the capillary by a pressure gradient. The exit side of the capillary caries a potential of 100-200 V.[226] When ions exit the capillary they are guided directly into the ion funnels described earlier.

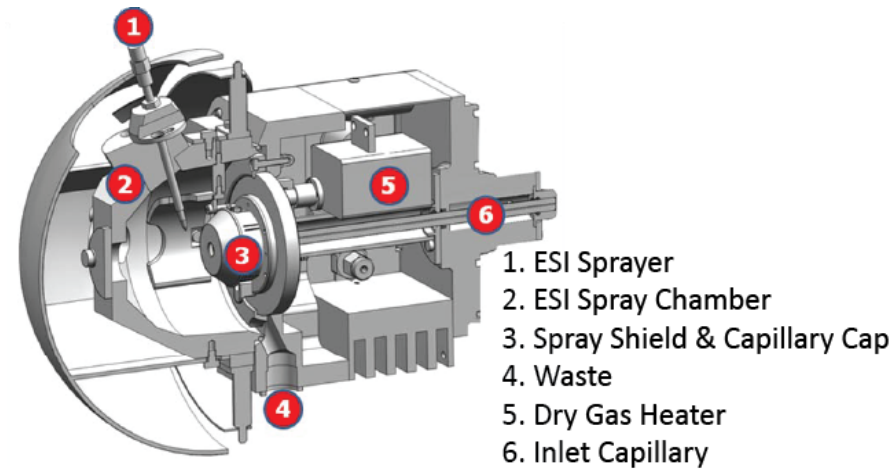


Figure 2-6. *A schematic overview of the Bruker Apollo ESI Source. Adapted from*[233].

For our experiments, the sample concentration used for ESI is around 10^{-6} mol/L. Typical solvents used in ESI are methanol, acetonitrile and water, usually in a mixture. In the experiments described in this thesis solutions of 50:50

acetonitrile/water were used and about 0.1% formic acid was added to enhance the formation of protonated peptides. For doubly protonated peptides about 0.5% formic acid was added to the solution. During the experiments the flow rate of the syringe pump was set at 120-180 µL/hr.

2.2.2. The free electron laser FELIX

The Free Electron Laser for Infrared eXperiments (FELIX) is located in the FELIX Laboratory which is part of the Radboud University in Nijmegen. The FELIX laboratory houses besides the FELIX laser also two other lasers: the Free electron Laser for Advanced spectroscopy and high Resolution Experiments (FLARE) and the Free Electron Laser for Intra-Cavity Experiments (FELICE). FELIX comprises two lasers, FELIX-1 and FELIX-2, each working at different electron beam energies and producing radiation in complementary ranges of the IR spectrum. The two FELIX lasers together cover the wavelength range from 3 to 150 µm. An overview of the lasers is presented in Figure 2-7.

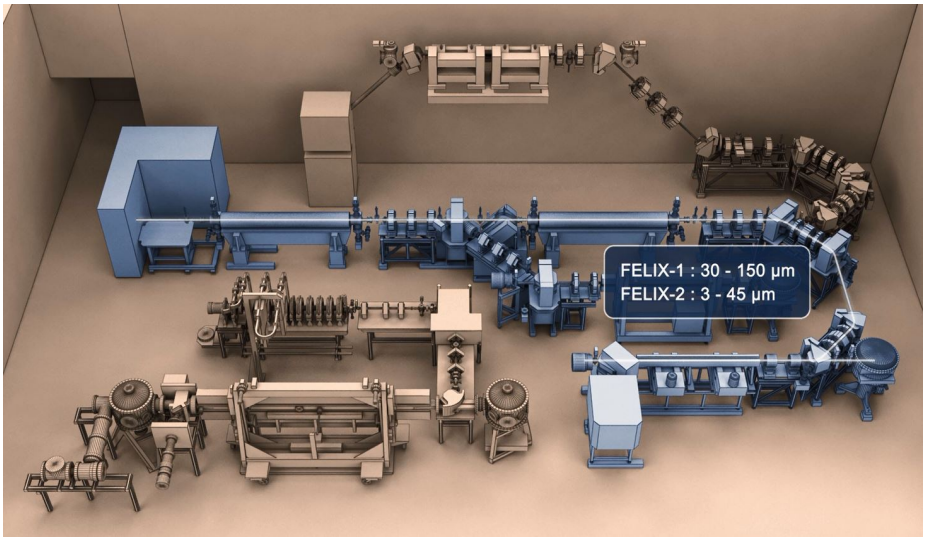


Figure 2-7. *The FELIX laboratory houses more lasers than the FELIX lasers. In this figure FELIX lasers 1 and 2 are highlighted. In this thesis laser light from FELIX-2 has been used for measuring IR spectra in the 600-2000 cm^{-1} region. Adapted from* [244].

A free electron laser distinguishes itself from other lasers by the fact that radiation is produced as a result of oscillating free electrons. Synchrotron radiation of wavelengths that can be varied over a wide range in the electromagnetic spectrum is produced by the accelerated electrons, essentially due to the transparent nature

of free, unbound electrons. The operational principle of the laser is based on the emission of radiation by relativistic electrons passing through an array of magnets of alternating polarity, the undulator. Because of the periodic acceleration, the charged particles (electrons) emit radiation at a wavelength $\lambda_{emitted}$ which is dependent on the undulator period λ_u, the electron energy (represented by Lorentz factor γ in equation 4) and the magnetic field strength (i.e. K value in equation [2-4]).

$$\lambda_{emitted} = \lambda_u \frac{1+K^2}{2\gamma^2} \quad [2\text{-}4]$$

With

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \quad [2\text{-}5]$$

In equation [2-5], the Lorentz factor γ is expressed as a function of the velocity (v) of the electrons and the speed of light (c). By varying the λ_u, γ and K parameters, a FEL can in principle produce light at any desired wavelength.

FELIX basically consists of an electron gun, where free electrons are generated, an accelerator where the electrons are accelerated (to 30 – 50 MeV), the undulator, and an optical cavity set up by two mirrors. A schematic overview of the components of FELIX is shown in Figure 2-8. The FELIX electron accelerator consists of two RF linear accelerators (LINAC 1 and LINAC 2) each accelerating the electrons to different energies. When the electrons reach the required energy, the electron beam is bent into the undulator which consists of two rows of samarium-cobalt permanent magnets that induce the oscillatory motion of the electrons. The tunability of FELIX is achieved by changing the distance between the two arrays of magnets using a stepper motor which changes the magnetic field strength (parameter K) in equation 4.[181,243] A small fraction of the light produced is coupled out of the cavity through a small hole in one of the mirrors and is transported through a beam transport system to the experimental set up, in this case the Bruker ion trap. Infrared transparent windows separate the beamline from the experimental set up. To avoid absorptions by atmospheric gases the beam line is evacuated.

The FELIX laser is a pulsed laser source. The electron beam and optical beam consists of macropulses, that consist of a train of micropulses. The duration of a FELIX macropulse is 5-10 μs with a spacing of 100 ms between two macropulses.

Micropulses are spaced by 1 ns in our experiments. This is illustrated in Figure 2-9. FELIX pulse energies can go up to about 100 mJ per macropulse corresponding to a micropulse peak energy of about 5 µJ. The general specifications of FELIX are summarized in Table 2-2.

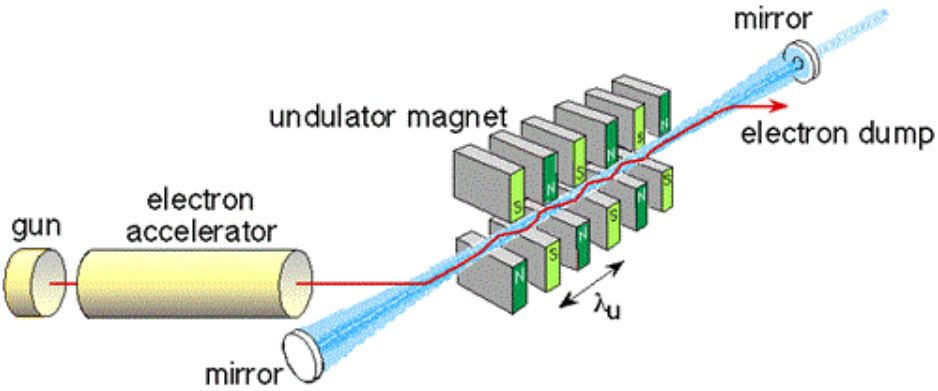


Figure 2-8. *The basic components of the free-electron laser FELIX: an electron accelerator to provide a beam of relativistic electrons, an undulator in which the electrons perform a transverse oscillation and emit radiation, and two mirrors forming an optical resonator in which the radiation is amplified by stimulated emission. Adapted from*[244].

In this thesis, experiments have been carried out in the 5.0 – 14.0 µm region of the IR spectrum using the FELIX-2 beam line. The bandwidth is Fourier-transform limited by the micropulse duration and was set at about 0.5% of the central frequency in our experiments which is about 3 - 10 cm^{-1} (corresponding to a micropulse duration in the order of 1-6 ps). Both bandwidth and central wavelength are measured online using a grating spectrometer equipped with an array detector having a dispersion of 0.02 µm per pixel. The accuracy of the wavelength calibration is about ± 0.01 µm. The difference between the set wavelength and actual wavelength measured by the spectrometer is fitted with a 3^{rd} order polynomial and subtracted from the saved set wavelength. Ions were usually irradiated with two FELIX macropulses with a repetition rate typically set at 10Hz. In our experiments a laser pulse energy of 5 - 60 mJ was used. Attenuation of the FELIX beam power is used in order to not completely deplete the signal of the precursor ion during the measurements. The power as function of wavelength is measured with a pyroelectric energy sensor and fitted with a 3^{rd} order polynomial to generate a power curve. Although IRMPD can have a non-linear response to laser power fluctuations, we use the power curve to linearly correct the fragment

ion yield. Severe deviations from linearity need to be characterized by recording the spectrum at different laser power levels.

	FELIX	OPO
Frequency (cm^{-1})	66-3700	2100-4000
Macropulse Energy (mJ)	< 150	2 – 20
M-pulse rep rate (Hz)	≤ 10	10
Bandwidth	0.2-5%	3 cm^{-1}

Table 2-2. *Specifications of the used lasers* [244].

2.2.3. Optical parametric oscillator (OPO-laser)

Measuring the IRMPD spectrum in the FELIX range is not always sufficient to determine the molecular structure of an ion. Therefore IRMPD spectra are measured in the 3000-3800 cm^{-1} region in addition to the spectra measured with FELIX. Characteristic absorptions in the 3000-3800 region of the infrared spectrum are O-H stretches in alcohol and carboxylic acid groups and N-H stretches of the N-terminal amine, the amide linkages and of some residue side chains (see table 2-1). Laser light in this region of the electromagnetic spectrum can be generated by FELIX lasing in third harmonic (with a pulse energy up to 10 mJ) or by a table top Optical Parametric Oscillator (OPO)-laser (with a pulse energy up to 20 mJ).

The OPO/OPA (LaserVision, Bellevue, WA, USA) at the FELIX laboratory is designed to convert the fixed frequency radiation of a 1064 nm Nd:YAG pump laser source (Innolas Spitlight 600, Munich, Germany) into a tunable output beam in the mid-IR range with a bandwidth of approximately 3 cm^{-1}.

Tunable infrared radiation from the OPO-laser is generated by frequency down conversion. The 1064 nm output of a pulsed Nd:YAG laser is split and injected into an optical parametric amplifier (OPA stage) and an OPO oscillator (OPO stage). In the OPO stage the frequency of the beam is doubled to 532 nm by a KTP crystal. From here a single pump laser photon is split into two new photons of lower frequency: the signal (around 800 nm) and the idler (around 1600 nm). The wavelengths of these complementary photons are changed by changing the angle of the birefringent KTP crystal, thereby changing the wavelengths at which phase-matching occurs. For amplification of the beams the oscillator crystal is placed inside a resonator cavity. Next, the idler beam is mixed with 1064 nm light in the OPA stage. Difference frequency generation in four KTA crystals generates light in the 3 μm spectral range. The remaining 1064 nm light is removed using a dichroic filter and the signal beam around 3 μm is filtered out using a polarizer.[181,245]

The wavelength range employed in the experiments described in this thesis is from 2.6-3.5 μm. The mid IR frequency of the OPA is calibrated by measuring the frequency of the Nd:YAG pump laser and signal beam of the OPO stage with a wavemeter (HighFinesse WS5). By subtracting the frequency of the pump laser from the frequency of the light that leaves as the signal beam of the OPO stage the actual frequency of the 3 μm light is determined.

2.3. Theoretical calculations

To extract detailed molecular structure information from the measured IRMPD spectra of the peptide fragment ions, we compare them with theoretically predicted spectra. We thereby take into account that they may be in one of many different isomeric, protomeric and conformeric forms. The actual structure of the species present in the mass spectrometer can then be assigned on the basis of a match between experimental and theoretical spectra. We employ a combination of Density Functional Theory (DFT) and Molecular Mechanics / Molecular Dynamics (MM/MD). DFT is applied for geometry optimization and prediction of the IR spectrum. Prior to DFT optimization, the potential energy landscape was usually probed using MM/MD methods to locate different conformational forms of the system. These calculations were performed at the Lisa and Cartesius LINUX clusters available at the SurfSARA supercomputer center.

To computationally model and predict CID MS/MS spectra and to visualize the reaction mechanism during CID, chemical dynamics simulations were also performed, using the local cluster available at Sorbonne University in Paris. The measured spectra and derived structures can serve as stringent benchmarks for these novel type of theoretical investigations.

2.3.1. Density functional theory

To optimize the geometry of a candidate structure to its minimum-energy geometry and to calculate the vibrational frequencies, density functional theory (DFT) calculations have been employed. The efficiency of the DFT approach combined with its high, post Hartree-Fock level of accuracy makes these methods to be the best option for the study of the systems under investigation here, consisting of several dozens of atoms. In general, DFT is able to adequately describe also non-covalent intramolecular interactions of interest in the studied peptide ions, such as especially hydrogen-bonding interactions.

All DFT calculations are carried out with the Gaussian 09 software package.[246] The calculation optimizes the geometry in an iterative way, in which the energy and the energy gradient are computed at every step. The minimum-energy structure is considered to be found when the energy and Root Mean Square (RMS) forces are converged to within predetermined cut-off values. All calculations described in this thesis make use of the B3LYP functional [247-249], which has been shown to be very reliable for the prediction of harmonic vibrational frequencies.[216] The 6-31++G(d,p) and 6-311+G(d,p) basis sets are used as they are both computationally efficient and sufficiently large to give reliable predictions of the infrared spectra. These basis sets use Gaussian functions to construct the basis functions, with the core electrons being modeled by 6 Gaussians, and split-valence double-zeta (31) and triple-zeta (311) functions modeling the valence electrons. The "+" or "++" signs indicate the addition of diffuse functions allowing the electrons to occupy a larger amount of space around the nuclei. The "(d,p)" indicates the addition of polarization functions, specifically p-functions to s-orbitals (hydrogen atoms) and d-functions to p-orbitals (heavy atoms).[250] For the studied peptide-ions the dispersion forces are normally considered to be weak, however, for the radical z-type fragment ions which have aromatic residues, these dispersion forces may play a significant role in the peptide conformation.[251] Dispersion forces are usually ignored at the DFT level and special functionals including different methods to correct for dispersion have therefore been developed, such as B3LYP-D and M06-2X. These functionals are also employed to model the structure of the z-ions.

The calculated vibrational frequencies are convoluted with a Gaussian line shape function of 15 to 25 cm^{-1} full width at half maximum (FWHM) to facilitate comparison with experimental spectra. A linear scaling factor of 0.975 for vibrational frequencies in the 800-2000 cm^{-1} range is used, while vibrations in the 2800-3800 cm^{-1} range are scaled with a factor of 0.955. This is a standard method to empirically correct for deviations from the harmonic approximation used in the frequency calculations.

To more accurately evaluate the relative electronic energies of different isomers and conformers of the system under study, single-point energies have also been computed at the MP2 [250,252] level of theory, using the nuclear geometry optimized at the B3LYP level of theory.

2.3.2. Molecular mechanics and molecular dynamics

In order to explore the potential energy surface (PES) of a peptide or fragment structure and find the lowest-energy conformer, a molecular mechanics/molecular dynamics approach has been used. In the MM/MD approach, nuclear motion is described by the laws of classical mechanics instead of quantum mechanics to reduce computational time and quickly generate possible conformational geometries. The potential energy of the molecule follows from the construction of the force field and includes bond stretching, angle bending, torsion as well as non-covalent interactions of the atoms in the molecule.[250] Force fields are applied to calculate the potential energy and the forces from the nuclear coordinates. Different force fields (FF's) have been developed, each optimized for different classes of molecular systems. For the peptides studied in this thesis, the General AMBER Force Field (GAFF) as implemented in molecular mechanics package AMBER 12[253] is employed.[254]

During the Molecular Dynamics procedure, a simulated annealing approach is applied in which the molecule is heated up to a maximum temperature of 1000 K while its dynamics is classically modeled using the force field and a 1 fs step size. The increased temperature enables the molecule to overcome (nearly) all transition state barriers in the PES and adopt any conformational structure. Subsequently, the temperature is gradually decreased condensing the molecule into one of the low-energy potential wells of the PES defined by the force field; these structures are stored. When 500 structures have thus been recorded, we assume to have sampled the PES sufficiently and these structures are then grouped based on structural similarity by using a clustering script; the RMS atom position criteria are chosen such that this grouping results in 20 to 30 structures. These resulting structures are then used as input for structure optimization and vibrational frequency calculation at the DFT level using the Gaussian 09 program, as described above. Further information about this computational protocol is given in refs. [161,188,255,256].

2.3.3. Chemical dynamics simulations

Problems for protein sequencing algorithms arise when proteins undergo fragmentation reactions that deviate from what is expected. Examples are the presence of post translational modifications, the involvement of the amino acid side chain in the fragmentation mechanism, or the re-opening of cyclic intermediates which leads to a permutation of the original sequence after applying CID.[85,86,257] Theoretical mass spectrometry, i.e. predicting MS/MS spectra, holds

great promise for better and more refined algorithms that can eventually lead to improved peptide and protein sequencing.[258] Moreover, theoretical mass spectrometry can aid in the interpretation of MS/MS spectra in general, and help us to understand peptide fragmentation patterns and to develop mechanistic insights.

Competing decomposition pathways of a molecule can be theoretically modeled by quantum-chemical calculations to locate the minima and transition states on the PES in combination with Rice-Ramsperger-Kassel-Marcus (RRKM)[259-262] / Quasi-Equilibrium Theory (QET)[263] statistical theories to determine the dissociation rate constants as a function of the internal energy of the molecule[264]. Detailed descriptions of these theories can be found in refs. [80,265-267]. This approach supposes that the minimum energy pathways are relevant in fragmentation[80] and that all relevant transition states on the PES can be identified. Knowing the dissociation pathways and their relative branching ratios then allows to predict an MS/MS spectrum.

An alternative approach to identify reaction pathways and their branching ratios is the use of chemical dynamics simulations modeling the fragmentation of a peptide.[268] In these simulations, the atom positions of the system evolve classically on a quantum-mechanically determined PES, which, for reasons of computational efficiency, is calculated at each step (on-the-fly) for the actual geometry of the molecule. A large number of trajectories, with randomized initial conditions, is simulated for reasons of statistical sampling. High-level theoretical methods cannot be employed to calculate the PES due to the excessive computational cost that would be required to simulate the large number of trajectories. Semi-empirical Hamiltonians are therefore often used instead of more costly DFT methods. Furthermore, the length of the trajectories is usually limited to a few tens of picoseconds. Trajectories are analyzed for reactivity and for the elemental composition and structure of the resulting fragments; from the number of occurrences of dissociative pathways into specific m/z channels, a predicted CID MS/MS spectrum can be constructed.

Two types of ion activation can be employed in the simulations, single-collision activation and internal energy activation.[258] In the first, a collision between the peptide and an inert gas atom with defined kinetic energy and randomized impact parameter energizes the molecule setting off the trajectory. In the second activation method, a fixed amount of internal energy is given to the ion that is randomly distributed among its vibrational modes without modeling the actual

collision. In this approach the (dissociative) trajectories can be analyzed as a function of internal energy.[258] The chemical dynamics code used in this thesis is VENUS[269], developed by Hase and coworkers, which is coupled with MOPAC[270] to simulate the collisions.

Bibliography

[1] Lehninger, A.L., Cox, M.M., Nelson, D.L., Principles of Biochemistry, 5th ed., W.H. Freeman & Co, New York, 2008.
[2] van Holde, K.E., Johnson, W.C., Ho, P.S., Principles of Physical Biochemistry, 2nd ed., Pearson Education, Upper Saddle River, New Jersey, 2005.
[3] Dale, J.W., von Schantz, M., From Genes to Genomes: Concepts and Applications of DNA Technology, 2nd ed., John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, 2008.
[4] Transcribe And Translate DNA Gene Sequences, https://cellcode.us/quotes/transcribe-and-translate-dna-gene-sequences.html
[5] Protein Synthesis-Translation, https://sites.google.com/site/proteinsynthesistranslation/home/terms
[6] Xie, H., Vucetic, S., Iakoucheva, L.M., Oldfield, C.J., Dunker, A.K., Obradovic, Z., Uversky, V.N., *J. Proteome Res.*, **6**, (2007), 1917-1932.
[7] Karve, T.M., Cheema, A.K., *J. Amino Acids*, **2011**, (2011), 1-13.
[8] Ohtsubo, K., Marth, J.D., *Cell*, **126**, (2006), 855-867.
[9] Jaeken, J., *Ann. N. Y. Acad. Sci.*, **1214**, (2010), 190-198.
[10] de Laat, P., Rodenburg, R., Smeitink, J., Mitochondrial Oxidative Phosphorylation Disorders, in: N. Blau, M. Duran, K.M. Gibson, C. Dionisi Vici (Eds.) Physician's Guide to the Diagnosis, Treatment, and Follow-Up of Inherited Metabolic Diseases, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 337-359.
[11] Speiser, P.W., White, P.C., *N. Engl. J. Med.*, **349**, (2003), 776-788.
[12] Santos, A.L., Lindner, A.B., *Oxid. Med. Cell Longev.*, **2017**, (2017), 1-19.
[13] Geiger, T., Clarke, S., *J. Biol. Chem.*, **262**, (1987), 785-794.
[14] Dunkelberger, E.B., Buchanan, L.E., Marek, P., Cao, P., Raleigh, D.P., Zanni, M.T., *J. Am. Chem. Soc.*, **134**, (2012), 12658-12667.
[15] Robinson, N.E., Robinson, A.B., *Proc. Natl. Acad. Sci. USA*, **98**, (2001), 944-949.
[16] Lindner, H., Helliger, W., *Exp. Geront.*, **36**, (2001), 1551-1563.
[17] Parker, M.W., *J. Biol. Phys.*, **29**, (2003), 341-362.
[18] Drenth, J., Principles of Protein X-ray Crystallography, 3 ed., Springer-Verlag, New York, 2007.
[19] Hinds, M.G., Norton, R.S., *Mol. Biotechnol.*, **7**, (1997), 315-331.
[20] Levitt, M.H., Spin Dynamics: Basics of Nuclear Magnetic Resonance, 2nd ed., John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, 2008.
[21] Frank, J., *Annu. Rev. Biophys. Biomol. Struct.*, **31**, (2002), 303-319.
[22] Jiménez, J.L., Guijarro, J.I., Orlova, E., Zurdo, J., Dobson, C.M., Sunde, M., Saibil, H.R., *EMBO J.*, **18**, (1999), 815-821.
[23] Zhang, W., Chipman, P.R., Corver, J., Johnson, P.R., Zhang, Y., Mukhopadhyay, S., Baker, T.S., Strauss, J.H., Rossmann, M.G., Kuhn, R.J., *Nat. Struct. Mol. Biol.*, **10**, (2003), 907.

[24] Borbat, P.P., McHaourab, H.S., Freed, J.H., *J. Am. Chem. Soc.*, **124**, (2002), 5304-5314.
[25] McHaourab, H.S., Steed, P.R., Kazmier, K., *Structure*, **19**, (2011), 1549-1561.
[26] Lippert, J.L., Tyminski, D., Desmeules, P.J., *J. Am. Chem. Soc.*, **98**, (1976), 7075-7080.
[27] Edman, P., *Acta Chem. Scand.*, **4**, (1950), 283-293.
[28] Yamashita, M., Fenn, J.B., *J. Phys. Chem.*, **88**, (1984), 4451-4459.
[29] Karas, M., Hillenkamp, F., *Anal. Chem.*, **60**, (1988), 2299-2301.
[30] Brodbelt, J.S., *Anal. Chem.*, **88**, (2016), 30-51.
[31] Zhurov, K.O., Fornelli, L., Wodrich, M.D., Laskay, U.A., Tsybin, Y.O., *Chem. Soc. Rev.*, **42**, (2013), 5014-5030.
[32] Cooper, H.J., Håkansson, K., Marshall, A.G., *Mass Spectrom. Rev.*, **24**, (2005), 201-222.
[33] Wu, C., Tran, J.C., Zamdborg, L., Durbin, K.R., Li, M., Ahlf, D.R., Early, B.P., Thomas, P.M., Sweedler, J.V., Kelleher, N.L., *Nat. Methods*, **9**, (2012), 822.
[34] Tran, J.C., Zamdborg, L., Ahlf, D.R., Lee, J.E., Catherman, A.D., Durbin, K.R., Tipton, J.D., Vellaichamy, A., Kellie, J.F., Li, M., Wu, C., Sweet, S.M.M., Early, B.P., Siuti, N., LeDuc, R.D., Compton, P.D., Thomas, P.M., Kelleher, N.L., *Nature*, **480**, (2011), 254.
[35] Seidler, J., Zinn, N., Boehm, M.E., Lehmann, W.D., *Proteomics*, **10**, (2010), 634-649.
[36] Perkins, D.N., Pappin, D.J.C., Creasy, D.M., Cottrell, J.S., *Electrophoresis*, **20**, (1999), 3551-3567.
[37] Yates, J.R., Eng, J.K., Clauser, K.R., Burlingame, A.L., *J. Am. Soc. Mass Spectrom.*, **7**, (1996), 1089-1098.
[38] Craig, R., Beavis, R.C., *Bioinformatics*, **20**, (2004), 1466-1467.
[39] Geer, L.Y., Markey, S.P., Kowalak, J.A., Wagner, L., Xu, M., Maynard, D., Yang, X., Shi, W., Bryant, S.H., *J. Proteome Res.*, **3**, (2004), 958-964.
[40] Dancik, V., Addona, T.A., Clauser, K.R., Vath, J.E., Pevzner, P.A., *J. Comput Biol.*, **6**, (1999), 327-342.
[41] Taylor, J.A., Johnson, R.S., *Anal. Chem.*, **73**, (2001), 2594-2604.
[42] Cox, J., Hubner, N.C., Mann, M., *J. Am. Soc. Mass Spectrom.*, **19**, (2008), 1813-1820.
[43] Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., Lajoie, G., *Rapid Commun. Mass Spectrom.*, **17**, (2003), 2337-2342.
[44] Frank, A., Pevzner, P., *Anal. Chem.*, **77**, (2005), 964-973.
[45] Fischer, B., Roth, V., Roos, F., Grossmann, J., Baginsky, S., Widmayer, P., Gruissem, W., Buhmann, J.M., *Anal. Chem.*, **77**, (2005), 7265-7273.
[46] Mo, L., Dutta, D., Wan, Y., Chen, T., *Anal. Chem.*, **79**, (2007), 4870-4878.
[47] Pan, C., Park, B.H., McDonald, W.H., Carey, P.A., Banfield, J.F., VerBerkmoes, N.C., Hettich, R.L., Samatova, N.F.A., *BMC Bioinf.*, **11**, (2010), 118.
[48] Savitski, M.M., Nielsen, M.L., Kjeldsen, F., Zubarev, R.A., *J. Proteome Res.*, **4**, (2005), 2348-2354.

[49] Pitzer, E., Masselot, A., Colinge, J., *Proteomics*, **7**, (2007), 3051-3054.
[50] Pevtsov, S., Fedulova, I., Mirzaei, H., Buck, C., Zhang, X., *J. Proteome Res.*, **5**, (2006), 3018-3028.
[51] Liska, A.J., Shevchenko, A., *Trends Anal. Chem.*, **22**, (2003), 291-298.
[52] Fernandez-de-Cossio, J., Gonzalez, J., Satomi, Y., Shima, T., Okumura, N., Besada, V., Betancourt, L., Padron, G., Shimonishi, Y., Takao, T., *Electrophoresis*, **21**, (2000), 1694-1699.
[53] Savitski, M.M., Nielsen, M.L., Zubarev, R.A., *Anal. Chem.*, **79**, (2007), 2296-2302.
[54] Wells, J.M., McLuckey, S.A., *Methods in Enzymol.*, **402**, (2005), 148-185.
[55] Bordas-Nagy, J., Jennings, K.R., *Int. J. Mass Spectrom. and Ion Process.*, **100**, (1990), 105-131.
[56] Williams, E.R., Furlong, J.J.P., McLafferty, F.W., *J. Am. Soc. Mass Spectrom.*, **1**, (1990), 288-294.
[57] Gauthier, J.W., Trautman, T.R., Jacobson, D.B., *Anal. Chim. Acta.*, **246**, (1991), 211-225.
[58] Zubarev, R.A., Kelleher, N.L., McLafferty, F.W., *J. Am. Chem. Soc.*, **120**, (1998), 3265-3266.
[59] Syka, J.E.P., Coon, J.J., Schroeder, M.J., Shabanowitz, J., Hunt, D.F., *Proc. Natl. Acad. Sci. USA*, **101**, (2004), 9528-9533.
[60] Little, D.P., Speir, J.P., Senko, M.W., O'Connor, P.B., McLafferty, F.W., *Anal. Chem.*, **66**, (1994), 2809-2815.
[61] Schnier, P.D., Price, W.D., Jockusch, R.A., Williams, E.R., *J. Am. Chem. Soc.*, **118**, (1996), 7178-7189.
[62] Ly, T., Julian, R.R., *Angew. Chem. Int. Ed.*, **48**, (2009), 7130-7137.
[63] Roepstorff, P., Fohlman, J., *Biomed. Mass Spectrom.*, **11**, (1984), 601-601.
[64] Biemann, K., *Methods Enzymol.*, **193**, (1990), 886-887.
[65] Chu, I.K., Siu, C.-K., Lau, J.K.-C., Tang, W.K., Mu, X., Lai, C.K., Guo, X., Wang, X., Li, N., Xia, Y., Kong, X., Oh, H.B., Ryzhov, V., Tureček, F., Hopkinson, A.C., Siu, K.W.M., *Int. J. Mass Spectrom* **390**, (2015), 24-27.
[66] Paizs, B., Suhai, S., *Mass Spectrom. Rev.*, **24**, (2004), 508-524.
[67] Wysocki, V.H., Tsaprailis, G., Smith, L.L., Breci, L.A., *J. Mass Spectrom.*, **35**, (2000), 1399-1406.
[68] Boyd, R., Somogyi, A., *J. Am. Soc. Mass Spectrom.*, **21**, (2010), 1275-1278.
[69] Harrison, A.G., Yalcin, T., *Int. J. Mass Spectrom. and Ion Process.*, **165/166**, (1997), 339-347.
[70] Dongre, A.R., Jones, J.L., Somogyi, A., Wysocki, V.H., *J. Am. Chem. Soc.*, **118**, (1996), 8365-8374.
[71] Wu, R., McMahon, T.B., *Chem. Phys. Chem.*, **9**, (2008), 2826-2835.
[72] Gu, C., Somogyi, Á., Wysocki, V.H., Medzihradszky, K.F., *Anal. Chim. Acta.*, **397**, (1999), 247-256.
[73] Komáromi, I., Somogyi, Á., Wysocki, V.H., *Int. J. of Mass Spectrom.*, **241**, (2005), 315-323.

[74] Polfer, N.C., Oomens, J., Suhai, S., Paizs, B., *J. Am. Chem. Soc*, **129**, (2007), 5887-5897.
[75] Cox, K.A., Gaskell, S.J., Morris, M., Whiting, A., *J. Am. Chem. Soc.*, **7**, (1996), 522-531.
[76] Jones, J.L., Dongre, A.R., Somogyi, A., Wysocki, V.H., *J. Am. Chem. Soc.*, **116**, (1994), 8368-8369.
[77] Tang, X.J., Thibault, P., Boyd, R.K., *Anal. Chem.*, **65**, (1993), 2824-2834.
[78] Kim, M.-S., Pandey, A., *Proteomics*, **12**, (2012), 530-542.
[79] Qi, Y., Volmer, D.A., *Mass Spectrom. Rev.*, **36**, (2015), 4-15.
[80] Song, K., Spezia, R., Theoretical Mass Spectrometry: Tracing Ions with Classical Trajectories, 1 ed., Walter de Gruyter GmbH & Co KG, Berlin, 2018.
[81] Sleno, L., Volmer, D.A., *J. Mass Spectrom.*, **39**, (2004), 1091-1112.
[82] McLuckey, S.A., *J. Am. Soc. Mass Spectrom.*, **3**, (1992), 599.
[83] Shukla, A.K., Futrell, J.H., *J. Mass Spectrom.*, **35**, (2000), 1069.
[84] Jennings, K.R., *Int. J. Mass Spectrom.*, **200**, (2000), 479.
[85] Harrison, A.G., *J. Am. Soc. Mass Spectrom.*, **19**, (2008), 1776-1780.
[86] Harrison, A.G., Young, A.B., Bleiholder, C., Suhai, S., Paizs, B., *J. Am. Chem. Soc.*, **128**, (2006), 10364-10365.
[87] Wiesner, J., Premsler, T., Sickmann, A., *Proteomics*, **8**, (2008), 4466-4483.
[88] Mirgorodskaya, E., Roepstorff, P., Zubarev, R.A., *Anal. Chem.*, **71**, (1999), 4431-4436.
[89] Stensballe, A., Jensen, O.N., Olsen, J.V., Haselmann, K.F., Zubarev, R.A., *Rapid Commun. Mass Spectrom.*, **14**, (2000), 1793-1800.
[90] Håkansson, K., Cooper, H.J., Emmett, M.R., Costello, C.E., Marshall, A.G., Nilsson, C.L., *Anal. Chem.*, **73**, (2001), 4530-4536.
[91] Mikesh, L.M., Ueberheide, B., Chi, A., Coon, J.J., Syka, J.E.P., Shabanowitz, J., Hunt, D.F., *Biochim. Biophys. Acta Proteins Proteomics*, **1764**, (2006), 1811-1822.
[92] Yalcin, T., Harrison, A.G., *J. Mass Spectrom.*, **31**, (1996), 1237-1243.
[93] Ballard, K.D., Gaskell, S.J., *Int. J. Mass Spectrom. and Ion Process.*, **111**, (1991), 173-189.
[94] Harrison, A.G., *Mass Spectrom. Rev.*, **28**, (2009), 640-654.
[95] Paizs, B., Suhai, S., *Rapid Commun. Mass Spectrom.*, **16**, (2002), 375-389.
[96] Polce, M.J., Ren, D., Wesdemiotis, C., *J. Mass Spectrom.*, **35**, (2000), 1391-1398.
[97] Yalcin, T., Khouw, C., Csizmadia, I.G., Peterson, M.R., Harrison, A.G., *J. Am. Soc. Mass Spectrom.*, **6**, (1995), 1165-1174.
[98] Nold, M.J., Wesdemiotis, C., Yalcin, T., Harrison, A.G., *Int. J. Mass Spectrom. and Ion Process.*, **164**, (1997), 137-153.
[99] Nold, M.J., Cerda, B.A., Wesdemiotis, C., *J. Am. Soc. Mass Spectrom.*, **10**, (1999), 1-8.
[100] Paizs, B., Suhai, S., *Rapid Commun. Mass Spectrom.*, **15**, (2001), 2307-2323.
[101] Cordero, M.M., Houser, J.J., Wesdemiotis, C., *Anal. Chem.*, **65**, (1993), 1594-1601.

[102] Grzetic, J., Oomens, J., *Int. J. Mass Spectrom*, **354-355**, (2013), 70-77.
[103] Farrugia, J.M., O'Hair, R.A.J., Reid, G.E., *Int. J. Mass Spectrom.*, **210-211**, (2001), 71-87.
[104] Harrison, A.G., *J. Mass Spectrom.*, **38**, (2003), 174-187.
[105] Jonsson, A.P., Bergman, T., Jörnvall, H., Griffiths, W.J., *Rapid Commun. Mass Spectrom.*, **15**, (2001), 713-720.
[106] Paizs, B., Szlávik, Z., Lendvay, G., Vékey, K., Suhai, S., *Rapid Commun. Mass Spectrom.*, **14**, (2000), 746-755.
[107] Bythell, B.J., Maître, P., Paizs, B., *J. Am. Chem. Soc.*, **132**, (2010), 14766-14779.
[108] Allen, J.M., Racine, A.H., Berman, A.M., Johnson, J.S., Bythell, B.J., Paizs, B., Glish, G.L., *J. Am. Soc. Mass Spectrom.*, **19**, (2008), 1764-1770.
[109] Mueller, D.R., Eckersley, M., Richter, W.J., *Org. Mass Spectrom.*, **23**, (1988), 217-222.
[110] Salek, M., Lehmann, W.D., *J. Mass Spectrom.*, **38**, (2003), 1143-1149.
[111] Martin, D.B., Eng, J.K., Nesvizhskii, A.I., Gemmill, A., Aebersold, R., *Anal. Chem.*, **77**, (2005), 4870-4882.
[112] Ballard, K.D., Gaskell, S.J., *J. Am. Soc. Mass Spectrom.*, **4**, (1993), 477-481.
[113] Balta, B., Aviyente, V., Lifshitz, C., *J. Am. Soc. Mass Spectrom.*, **14**, (2003), 1192-1203.
[114] Reid, G.E., Simpson, R.J., O'Hair, R.A.J., *Int. J. Mass Spectrom.*, **190-191**, (1999), 209-230.
[115] Reid, G.E., Simpson, R.J., O'Hair, R.A.J., *J. Am. Soc. Mass Spectrom.*, **11**, (2000), 1047-1060.
[116] Csonka, I.P., Paizs, B., Lendvay, G., Suhai, S., *Rapid Commun. Mass Spectrom.*, **14**, (2000), 417-431.
[117] Dookeran, N.N., Yalcin, T., Harrison, A.G., *J. Mass Spectrom.*, **31**, (1996), 500-508.
[118] Hunt, D.F., Shabanowitz, J., Bai, D.L., *J. Am. Soc. Mass Spectrom.*, **26**, (2015), 1256-1258.
[119] Coon, J.J., Syka, J.E.P., Schwartz, J.C., Shabanowitz, J., Hunt, D.F., *Int. J. Mass Spectrom.*, **236**, (2004), 33-42.
[120] Compton, P.D., Strukl, J.V., Bai, D.L., Shabanowitz, J., Hunt, D.F., *Anal. Chem.*, **84**, (2012), 1781-1785.
[121] Gunawardena, H.P., He, M., Chrisman, P.A., Pitteri, S.J., Hogan, J.M., Hodges, B.D.M., McLuckey, S.A., *J. Am. Chem. Soc.*, **127**, (2005), 12627-12639.
[122] Coon, J.J., Ueberheide, B., Syka, J.E.P., Dryhurst, D.D., Ausio, J., Shabanowitz, J., Hunt, D.F., *Proc. Natl. Acad. Sci. USA*, **102**, (2005), 9463-9468.
[123] McAlister, G.C., Berggren, W.T., Griep-Raming, J., Horning, S., Makarov, A., Phanstiel, D., Stafford, G., Swaney, D.L., Syka, J.E.P., Zabrouskov, V., Coon, J.J., *J. Proteome Res.*, **7**, (2008), 3127-3136.
[124] Riley, N.M., Coon, J.J., *Anal. Chem.*, **90**, (2018), 40-64.
[125] Tureček, F., Julian, R.R., *Chem. Rev.*, **113**, (2013), 6691-6733.

[126] Guan, Z., Kelleher, N.L., O'Connor, P.B., Aaserud, D.J., Little, D.P., McLafferty, F.W., *Int. J. Mass Spectrom. and Ion Process.*, **157-158**, (1996), 357-364.
[127] Stephenson, J.L., McLuckey, S.A., *Int. J. Mass Spectrom. and Ion Process.*, **162**, (1997), 89-106.
[128] Zubarev, R.A., *Mass Spectrom. Rev.*, **22**, (2003), 57-77.
[129] Tsaprailis, G., Nair, H., Somogyi, Á., Wysocki, V.H., Zhong, W., Futrell, J.H., Summerfield, S.G., Gaskell, S.J., *J. Am. Chem. Soc.*, **121**, (1999), 5142-5154.
[130] Turecek, F., *J. Am. Chem. Soc.*, **125**, (2003), 5954-5963.
[131] Laskin, J., Yang, Z., Lam, C., Chu, I.K., *Anal. Chem.*, **79**, (2007), 6607-6614.
[132] Tsybin, Y.O., Fornelli, L., Stoermer, C., Luebeck, M., Parra, J., Nallet, S., Wurm, F.M., Hartmer, R., *Anal. Chem.*, **83**, (2011), 8919-8927.
[133] Breuker, K., Oh, H., Horn, D.M., Cerda, B.A., McLafferty, F.W., *J. Am. Chem. Soc.*, **124**, (2002), 6407-6420.
[134] Chi, A., Huttenhower, C., Geer, L.Y., Coon, J.J., Syka, J.E.P., Bai, D.L., Shabanowitz, J., Burke, D.J., Troyanskaya, O.G., Hunt, D.F., *Proc. Natl. Acad. Sci. USA*, **104**, (2007), 2193-2198.
[135] Good, D.M., Wirtala, M., McAlister, G.C., Coon, J.J., *Mol. Cell. Proteomics.*, **6**, (2007), 1942-1951.
[136] Hurtado, P.P., O'Connor, P.B., *Mass Spectrom. Rev.*, **2**, (2012), 609-625.
[137] Kjeldsen, F., Zubarev, R.A., *J. Am. Chem. Soc.*, **125**, (2003), 6628-6629.
[138] Falth, M., Savitski, M.M., Nielsen, M.L., Kjeldsen, F., Andren, P.E., Zubarev, R.A., *Anal. Chem.*, **80**, (2008), 8089-8094.
[139] Wee, S., O'Hair, R.A.J., McFadyen, W.D., *Rapid Commun. Mass Spectrom.*, **16**, (2002), 884-890.
[140] Johnson, R.S., Martin, S.A., Biemann, K., Stults, J.T., Watson, J.T., *Anal. Chem.*, **59**, (1987), 2621-2625.
[141] Kruger, N.A., Zubarev, R.A., Horn, D.M., McLafferty, F.W., *Int. J. Mass Spectrom.*, **185**, (1999), 787-793.
[142] Turecek, F., Syrstad, E.A., *J. Am. Chem. Soc.*, **125**, (2003), 3353-3369.
[143] Syrstad, E.A., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **16**, (2005), 208-224.
[144] Sawicka, A., Skurski, P., Hudgins, R.R., Simons, J., *J. Phys. Chem. B.*, **107**, (2003), 13505-13511.
[145] Syrstad, E.A., Tureček, F., *J. Phys. Chem. A.*, **105**, (2001), 11144-11155.
[146] Zubarev, R.A., Horn, D.M., Fridriksson, E.K., Kelleher, N.L., Kruger, N.A., Lewis, M.A., Carpenter, B.K., McLafferty, F.W., *Anal. Chem.*, **72**, (2000), 563-573.
[147] Anusiewicz, I., Berdys-Kochanska, J., Simons, J., *J. Phys. Chem. A.*, **109**, (2005), 5801-5813.
[148] Anusiewicz, I., Berdys-Kochanska, J., Skurski, P., Simons, J., *J. Chem. Phys. A.*, **110**, (2006), 1261-1266.
[149] Sobczyk, M., Simons, J., *J. Phys. Chem. A.*, **110**, (2006), 7519-7527.
[150] Chen, X., Tureček, F., *J. Am. Chem. Soc.*, **128**, (2006), 12520-12530.
[151] Simons, J., *Chem. Phys. Lett.*, **484**, (2010), 81-95.

[152] Patriksson, A., Adams, C., Kjeldsen, F., Raber, J., van der Spoel, D., Zubarev, R.A., *Int. J. of Mass Spectrom.*, **248**, (2006), 124-135.
[153] Wodrich, M.D., Zhurov, K.O., Vorobyev, A., Hamidane, H.B., Corminboeuf, C., Tsybin, Y.O., *J. Phys. Chem. B.*, **116**, (2012), 10807-10815.
[154] Wodrich, M.D., Zhurov, K.O., Corminboeuf, C., Tsybin, Y.O., *J. Phys. Chem. B.*, **118**, (2014), 2985-2992.
[155] Zubarev, R.A., Haselmann, K.F., Budnik, B., Kjeldsen, F., Jensen, F., *Eur. J. Mass Spectrom.*, **8**, (2002), 337-349.
[156] Zubarev, R.A., Kruger, N.A., Fridriksson, E.K., Lewis, M.A., Horn, D.M., Carpenter, B.K., McLafferty, F.W., *J. Am. Chem. Soc.*, **121**, (1999), 2857-2862.
[157] Świerszcz, I., Skurski, P., Simons, J., *J. Phys. Chem. A.*, **116**, (2012), 1828-1837.
[158] Tureček, F., Chen, X., Hao, C., *J. Am. Chem. Soc.*, **130**, (2008), 8818-8833.
[159] Tureček, F., Syrstad, E.A., Seymour, J.L., Chen, X., Yao, C., *J. Mass Spectrom.*, **38**, (2003), 1093-1104.
[160] Anusiewicz, I., Skurski, P., Simons, J., *J. Phys. Chem. B.* , **118**, (2014), 7892-7901.
[161] Martens, J., Grzetic, J., Berden, G., Oomens, J., *Nat. Comm.*, **7**, (2016), 11754.
[162] Moore, B.N., Ly, T., Julian, R.R., *J. Am. Chem. Soc.*, **133**, (2011), 6997-7006.
[163] Siu, C., Ke, Y., Orlova, G., Hopkinson, A.C., Siu, K.W.M., *J. Am. Soc. Mass Spectrom.*, **19**, (2008), 1799-1807.
[164] Viehe, H.G., Janousek, Z., Merenyi, R., Stella, L., *Acc. Chem. Res.*, **18**, (1985), 148-154.
[165] Bagheri-Majdi, E., Ke, Y., Orlova, G., Chu, I.K., Hopkinson, A.C., Siu, K.W.M., *J. Phys. Chem. B.*, **108**, (2004), 11170-11181.
[166] Zhao, J., Siu, K.W.M., Hopkinson, A.C., *Phys. Chem. Chem. Phys.*, **10**, (2008), 281-288.
[167] Croft, A.K., Easton, C.J., Radom, L., *J. Am. Chem. Soc.*, **125**, (2003), 4119-4124.
[168] Burgers, P.C., Holmes, J.L., Terlouw, J.K., van Baar, B., *Org. Mass Spectrom.*, **20**, (1985), 202-206.
[169] Cooper, H.J., Hudgins, R.R., Håkansson, K., Marshall, A.G., *J. Am. Soc. Mass Spectrom.*, **13**, (2002), 241-249.
[170] Leymarie, N., Costello, C.E., O'Connor, P.B., *J. Am. Chem. Soc.*, **125**, (2003), 8949-8958.
[171] Haselmann, K.F., Budnik, B.A., Kjeldsen, F., *Eur. J. Mass Spectrom.*, **8**, (2002), 461-469.
[172] Cooper, H.J., Hudgins, R.R., Håkansson, K., Marshall, A.G., *Int. J. Mass Spectrom.*, **228**, (2003), 723-728.
[173] Xia, Q., Lee, M.V., Rose, C.M., Marsh, A.J., Hubler, S.L., Wenger, C.D., Coon, J.J., *J. Am. Soc. Mass Spectrom.*, **22**, (2011), 255-264.
[174] Asakawa, D., Wada, Y., *J. Phys. Chem. B.*, **118**, (2014), 12318-12325.
[175] Sun, Q., Nelson, H., Ly, T., Stoltz, B.M., Julian, R.R., *J. Proteome Res.*, **8**, (2009), 958-966.

[176] Li, X., Lin, C., Han, L., Costello, C.E., O'Connor, P.B., *J. Am. Soc. Mass Spectrom.*, **21**, (2010), 646-656.
[177] Kjeldsen, F., Haselmann, K.F., Budnik, B.A., Jensen, F., Zubarev, R.A., *Chem. Phys. Lett.*, **356**, (2002), 201-206.
[178] Kanu, A.B., Dwivedi, P., Tam, M., Matz, L., Hill, H.H., *J. Mass Spectrom.*, **43**, (2008), 1-22.
[179] Weis, D.D., Wales, T.E., Engen, J.R., Hotchko, M., Ten Eyck, L.F., *J. Am. Soc. Mass Spectrom.*, **17**, (2006), 1498-1509.
[180] Zhang, H., Li, X.-J., Martin, D.B., Aebersold, R., *Nature Biotechnol.*, **21**, (2003), 660-666.
[181] Rijs, A.M., Oomens, J., Gas-Phase IR Spectroscopy and Structure of Biological Molecules, Springer, Switzerland, 2015.
[182] Lucas, B., Grégoire, G., Lemaire, J., Maître, P., Ortega, J., Rupenyan, A., Reimann, B., Schermann, J.P., Desfrançois, C., *Phys. Chem. Chem. Phys.*, **6**, (2004), 2659-2663.
[183] Heaton, A.L., Bowman, V.N., Oomens, J., Steill, J.D., Armentrout, P.B., *J. Phys. Chem. A.*, **113**, (2009), 5519-5530.
[184] Correia, C.F., Balaj, P.O., Scuderi, D., Maitre, P., Ohanessian, G., *J. Am. Chem. Soc.*, **130**, (2008), 3359-3370.
[185] Patrick, A.L., Stedwell, C.N., Schindler, B., Compagnon, I., Berden, G., Oomens, J., Polfer, N.C., *Int. J. Mass Spectrom*, **379**, (2015), 26-32.
[186] Bythell, B.J., Erlekam, U., Paizs, B., Maître, P., *Chem. Phys. Chem.*, **10**, (2009), 883-885.
[187] Polfer, N.C., Oomens, J., *Phys. Chem. Chem. Phys.*, **9**, (2007), 3804-3817.
[188] Martens, J., Grzetic, J., Berden, G., Oomens, J., *Int. J. of Mass Spectrom.*, **377**, (2015), 179-187.
[189] Frison, G., van der Rest, G., Tureček, F., Besson, T., Lemaire, J., Maître, P., Chamot-Rooke, J., *J. Am. Chem. Soc.*, **130**, (2008), 14916-14917.
[190] Oomens, J., Young, S., Molesworth, S., van Stipdonk, M., *J. Am. Soc. Mass Spectrom.*, **20**, (2009), 334-339.
[191] Yoon, S.H., Chamot-Rooke, J., Perkins, B.R., Hilderbrand, A.E., Poutsma, J.C., Wysocki, V.H., *J. Am. Chem. Soc.*, **130**, (2008), 17644-17645.
[192] Perkins, B.R., Chamot-Rooke, J., Yoon, S.H., Gucinski, A.C., Somogyi, A., Wysocki, V.H., *J. Am. Chem. Soc.*, **131**, (2009), 17528-17529.
[193] Erlekam, U., Bythell, B.J., Scuderi, D., van Stipdonk, M., Paizs, B., Maitre, P., *J. Am. Chem. Soc.*, **131**, (2009), 11503-11508.
[194] Piatkivskyi, A., Osburn, S., Jaderberg, K., Grzetic, J., Steill, J.D., Oomens, J., Zhao, J., Lau, J.K.-C., Verkerk, U.H., Hopkinson, A.C., Siu, K.W.M., Ryzhov, V., *J. Am. Soc. Mass Spectrom.*, **24**, (2013), 513-523.
[195] Osburn, S., Berden, G., Oomens, J., O'Hair, R.A.J., Ryzhov, V., *J. Am. Soc. Mass Spectrom.*, **23**, (2012), 1019-1023.
[196] Osburn, S., Berden, G., Oomens, J., O'Hair, R.A.J., Ryzhov, V., *J. Am. Soc. Mass Spectrom.*, **22**, (2011), 1794-1803.

[197] Steill, J.D., Zhao, J., Siu, C.-K., Ke, Y., Verkerk, U.H., Oomens, J., Dunbar, R.C., Hopkinson, A.C., Siu, K.W.M., *Angew. Chem. Int. Ed.*, **47**, (2008), 9666-9668.
[198] Fung, Y.M.E., Besson, T., Lemaire, J., Maître, P., Zubarev, R.A., *Angew. Chem. Int. Ed.*, **121**, (2009), 8490-8492.
[199] Sinha, R.K., Erlekam, U., Bythell, B.J., Paizs, B., Maitre, P., *J. Am. Soc. Mass Spectrom.*, **22**, (2011), 1645-1650.
[200] Polfer, N.C., Oomens, J., *Mass Spectrom. Rev.*, **28**, (2009), 468-494.
[201] Wu, R., McMahon, T.B., *J. Phys. Chem. B.*, **113**, (2009), 8767-8775.
[202] Scuderi, D., Correia, C.F., Balaj, O.P., Ohanessian, G., Lemaire, J., Maitre, P., *Chem. Phys. Chem.*, **10**, (2009), 1630-1641.
[203] Correia, C.F., Clavaguera, C., Erlekam, U., Scuderi, D., Ohanessian, G., *Chem. Phys. Chem.*, **9**, (2008), 2564-2573.
[204] Bythell, B.J., Dain, R.P., Curtice, S.S., Oomens, J., Steill, J.D., Groenewold, G.S., Paizs, B., Van Stipdonk, M.J., *J. Chem. Phys. A.*, **114**, (2010), 5076-5082.
[205] Chalifoux, A.M., Boles, G.C., Berden, G., Oomens, J., Armentrout, P.B., *Phys. Chem. Chem. Phys.*, **20**, (2018), 20712-20725.
[206] Oomens, J., von Helden, G., Meijer, G., *J. Phys. Chem. A.*, **108**, (2004), 8273-8278.
[207] Oomens, J., Polfer, N., Moore, D.T., van der Meer, L., Marshall, A.G., Eyler, J.R., Meijer, G., von Helden, G., *Phys. Chem. Chem. Phys.*, **7**, (2005), 1345-1348.
[208] Seo, J., Hoffmann, W., Warnke, S., Huang, X., Gewinner, S., Schöllkopf, W., Bowers, M.T., von Helden, G., Pagel, K., *Nat. Chem.*, **9**, (2016), 39-44.
[209] Chen, X., Yu, L., Steill, J.D., Oomens, J., Polfer, N.C., *J. Am. Chem. Soc.*, **131**, (2009), 18272-18282.
[210] Kempkes, L.J.M., Martens, J., Berden, G., Oomens, J., *J. Mass Spectrom.*, **53**, (2018), 1207-1213.
[211] Jašíková, L., Roithová, J., *Chem. Eur. J.;*, **24**, (2018), 3374-3390.
[212] Oomens, J., Polfer, N.C., Berden, G., Eyler, J.R., *Eur. J. Mass Spectrom.*, (2018).
[213] Baer, T., Dunbar, R.C., *J. Am. Soc. Mass Spectrom.*, **21**, (2010), 681-693.
[214] Verkerk, G., Broens, J.B., de Groot, P.A.M., Binas: informatieboek vwo-havo voor het onderwijs in de natuurwetenschappen, 5e druk, Wolters-Noordhoff, 2004.
[215] Valle, J.J., Eyler, J.R., Oomens, J., Moore, D.T., van der Meer, A.F.G., von Helden, G., Meijer, G., Hendrickson, C.L., Marshall, A.G., Blakney, G.T., *Rev. Sci. Instrum.*, **76**, (2005), 023103.
[216] Lemaire, J., Boissel, P., Heninger, M., Mauclaire, G., Bellec, G., Mestdagh, H., Simon, A., Caer, S.L., Ortega, J.M., Glotin, F., Maitre, P., *Phys. Rev. Lett.*, **89**, (2002), 273002.
[217] Brodbelt, J.S., Wilson, J.J., *Mass Spectrom. Rev.*, **28**, (2009), 390-424.
[218] Eyler, J.R., *Mass Spectrom. Rev.*, **28**, (2009), 448-467.
[219] Roithova, J., *Chem. Soc. Rev.*, **41**, (2012), 547-559.
[220] Cooper, D.E., Klimcak, C.M., Wessel, J.E., *Phys. Rev. Lett.*, **46**, (1981), 324-328.
[221] Bakker, J.M., Aleese, L.M., Meijer, G., von Helden, G., *Phys. Rev. Lett.*, **91**, (2003), 203003.

[222] Goebbert, D.J., Wende, T., Bergmann, R., Meijer, G., Asmis, K.R., *J. Phys. Chem. A.*, **113**, (2009), 5874-5880.
[223] Asvany, O., Brünken, S., Kluge, L., Schlemmer, S., *App. Phys. B.*, **114**, (2014), 203-211.
[224] Okumura, M., Yeh, L.I., Lee, Y.T., *J. Chem. Phys.*, **83**, (1985), 3705-3706.
[225] Chemistry LibreTexts Library - Bond Energies, https://chem.libretexts.org/Textbook_Maps/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Chemical_Bonding/Fundamentals_of_Chemical_Bonding/Bond_Energies
[226] Martens, J., Berden, G., Gebhardt, C.R., Oomens, J., *Rev. Sci. Instrum.*, **87**, (2016), 103108.
[227] Polfer, N.C., *Chem. Soc. Rev.*, **40**, (2011), 2211-2221.
[228] Oomens, J., Sartakov, B.G., Meijer, G., von Helden, G., *Int. J. Mass Spectrom.*, **254**, (2006), 1-19.
[229] von Helden, G., van Heijnsbergen, D., Meijer, G., *J. Phys. Chem. A.*, **107**, (2003), 1671-1688.
[230] Fridgen, T.D., *Mass Spectrom. Rev.*, **28**, (2009), 586-607.
[231] Joblin, C., Boissel, P., Leger, A., d'Hendecourt, L., Defourneau, D., *Astron. Astrophys.*, **299**, (1995), 835-846.
[232] Parneix, P., Basire, M., Calvo, F., *J. Phys. Chem. A*, **117**, (2013), 3954-3959.
[233] Bruker-Daltonics, AmaZon Series - User Manual, Bruker, Germany, 2009.
[234] March, R.E., Todd, J.F.J., Fundamentals of Ion Trap Mass Spectrometry, 1st ed., CRS, 1995.
[235] March, R.E., *J. Mass Spectrom.*, **32**, (1997), 351-369.
[236] Dass, C., Fundamentals of Contemporary Mass Spectrometry, 1st ed., Wiley, New Jersey, 2007.
[237] Paul, W., *Nobel Lecture*, (1989).
[238] March, R.E., *Int. J. Mass Spectrom.*, **200**, (2000), 285-312.
[239] Hartmer, R., Kaplan, D.A., Gebhardt, C.R., Ledertheil, T., Brekenfeld, A., *Int. J. Mass Spectrom.*, **276**, (2008), 82-90.
[240] Martens, J., Berden, G., Oomens, J., *Anal. Chem.*, **88**, (2016), 6126-6129.
[241] Fenn, J.B., *Nobel Lecture*, (2002).
[242] Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., Whitehouse, C.M., *Science*, **246**, (1989), 64-71.
[243] Oepts, D., van der Meer, A.F.G., van Amersfoort, P.W., *Infrared Phys. Technol.*, **36**, (1995), 297-308.
[244] FELIX-Laboratory, FELIX in more detail, https://www.ru.nl/felix/about-felix/about-felix/felix-laboratory/
[245] Bosenberg, W.R., Guyer, D.R., *J. Opt. Soc. Am. B.*, **10**, (1993), 1716-1722.
[246] Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G.A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H.P., Izmaylov, A.F., Bloino, J., Zheng, G., Sonnenberg, J.L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida,

M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J., J.A., Peralta, J.E., Ogliaro, F., Bearpark, M., Heyd, J.J., Brothers, E., Kudin, K.N., Staroverov, V.N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Rega, N., Millam, N.J., Klene, M., Knox, J.E., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Martin, R.L., Morokuma, K., Zakrzewski, V.G., Voth, G.A., Salvador, P., Dannenberg, J.J., Dapprich, S., Daniels, A.D., Farkas, O., Foresman, J.B., Ortiz, J.V., Cioslowski, J., Fox, D.J., *Gaussian, Inc., Wallingford, CT*, (2009).
[247] Becke, A.D., *J. Chem. Phys.*, **98**, (1993), 5648-5652.
[248] Miehlich, B., Savin, A., Stoll, H., Preuss, H., *Chem. Phys. Lett.*, **157**, (1989), 200-206.
[249] Becke, A.D., *Phys. Rev. A.*, **38**, (1988), 3098-3100.
[250] Cramer, C.J., Essentials of computational chemistry: theories and models, John Wiley & Sons, The Atrium, Southern Gate, Chichester, West Sussex, 2013.
[251] Jaeqx, S., PhD-thesis: Protein folding forces probed by infrared action spectroscopy supported by quantum chemical calculations, Radboud University, Nijmegen, 2014.
[252] Møller, C., Plesset, M.S., *Phys. Rev.*, **46**, (1934), 618-622.
[253] Case, D.A., Darden, T., Cheatham III, T.E., Simmerling, C., Wang, J., Duke, R.E., Luo, R., Walker, R.C., Zhang, W., Merz, K.M., Roberts, B.P., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Goetz, A.W., Kolossváry, I., Wong, K.F., Paesani, F., Vanicek, J., Wolf, R.M., Liu, J., Wu, X., Brozell, S.R., Steinbrecher, T., Gohlke, H., Cai Ye, Q.X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D.R., Mathews, D.H., Seetin, M.G., Salomon-Ferrer, R., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., Kollman, P.A., *University of California, San Francisco*, (2012).
[254] Salomon-Ferrer, R., Case, D.A., Walker, R.C., *Wiress. Comput. Mol. Sci.*, **3**, (2013), 198-210.
[255] Stedwell, C.N., Galindo, J.F., Gulyuz, K., Roitberg, A.E., Polfer, N.C., *J. Phys. Chem. A.*, **117**, (2013), 1181-1188.
[256] Grzetic, J., PhD-thesis: Chemistry of peptide fragmentation - Molecular structures by ion spectroscopy, Radboud University, Nijmegen, 2014.
[257] Tirado, M., Polfer, N.C., *Angew. Chem. Int. Ed.* , **51**, (2012), 6436-6438.
[258] Martin-Somer, A.M., J.; Grzetic, J.; Hase, W.L.; Oomens, J.; Spezia, R.;, *J. Phys. Chem. A*, (2018).
[259] Rice, H.C., *J. Am. Chem. Soc.*, **49**, (1927), 1617-1629.
[260] Kassel, L.S., *J. Phys. Chem.*, **32**, (1927), 225-242.
[261] Marcus, R.A., *J. Chem. Phys.*, **20**, (1952), 359-364.
[262] Marcus, R.A., *J. Phys. Chem.*, **55**, (1951), 894-908.
[263] Wallenstein, H.M., *Proc. Natl. Acad. Sci. U. S. A.*, **38**, (1952), 667-678.
[264] Baer, T., Hase, W.L., Unimolecular Reaction Dynamics. Theory and Experiments, Oxford University Press, New York, 1996.
[265] Di Giacomo, F., *J. Chem. Educ.* , **92**, (2015), 476-481.

[266] Bauer, C.A., Grimme, S., *J. Phys. Chem. A.*, **120**, (2016), 3755-3766.
[267] Armentrout, P.B., Ervin, K.M., Rodgers, M.T., *J. Phys. Chem. A.*, **112**, (2008), 10071-10085.
[268] Spezia, R., Martens, J., Oomens, J., Song, K., *Int. J. Mass Spectrom.*, **338**, (2015), 40-52.
[269] Hase, W.L., Duchovic, R.J., Hu, X., Komornicki, A., Lim, K.F., Lu, D.H., Peslherbe, G.H., Swamy, K.N., Vande-Linde, S.R., Zhu, L., Varandas, A., Wang, H., Wolf, R.J., *QCPE Bull.*, **16**, (1996), 671.
[270] Stewart, J.J.P., Fiedler, L.J., Zhang, P., Zheng, J., Rossi, I., Hu, W.-P., Lynch, G.C., Liu, Y.-P., Chuang, Y.-Y., Pu, J., Li, J., Cramer, C.J., Fast, P.L., Truhlar, D.G., *Department of Chemistry and Supercomputing Institute, University of Minnesota: Minneapolis, MN*, (2015).

Part II – Spectroscopic investigation of deamidation and dehydration reactions induced by collisional dissociation

Chapter 3

Deamidation reactions of protonated asparagine and glutamine investigated by ion spectroscopy

---

Deamidation of Asn and Gln residues is a primary route for spontaneous post-translational protein modification. Various structures have been proposed for the deamidation products of the protonated amino acids. Here we verify these structures by ion spectroscopy, as well as the structures of parallel and sequential fragmentation products. Infrared ion spectroscopy using the free electron laser FELIX has been applied to the reaction products from deamidation of protonated glutamine and asparagine in a tandem mass spectrometer. IR spectra were recorded over the 800-1900 cm^{-1} spectral range by infrared multiple-photon dissociation (IRMPD) spectroscopy. Molecular structures of the fragment ions are derived from comparison of the experimental spectra with spectra predicted for different candidate structures by density functional theory (DFT) calculations. $[Asn+H-NH_3]^+$ is found to possess a 3-amino succinic anhydride structure protonated on the amino group. The dissociation reaction involving loss of H_2O and CO forms a linear immonium ion. For $[Gln+H^+ - NH_3]^+$, the N-terminal nitrogen acts as the nucleophile leading to an oxo-proline product ion structure. For $[Gln+H^+ - NH_3]^+$ a sequential loss of $[CO + H_2O]$ is found, leading to a pyrolidone-like structure. We also confirm by IR spectroscopy that dehydration of protonated aspartic acid $[Asp+H]^+$ and glutamic acid $[Glu+H^+]$ leads to identical structures as to those found for the loss of NH_3 from $[Asn+H]^+$ and $[Gln+H]^+$. The structure determined for $[Asn+H]^+$ is in agreement with the suggested structure derived from measured and computed activation energies. Infrared ion spectra for the NH_3-loss product from $[Gln+H]^+$ establish that a different reaction mechanism occurs for this species as compared to $[Asn+H]^+$. For both amino acids, loss of NH_3 occurs from the side-chain[*].

[*] Adapted from: Kempkes, L. J. M.; Martens, J. K.; Grzetic, J.; Berden, G.; Oomens, J., Deamidation reactions of protonated asparagine and glutamine investigated by ion spectroscopy. *Rapid Commun. Mass Spectrom.* **2016,** *30*, 483-490.

3.1. Introduction

The amino acids glutamine (Gln, Q) and asparagine (Asn, N) feature an amide group in their side chain. Deamidation of Asn and Gln residues is a primary route for spontaneous post-translational protein modification[1]. It is a dominant pathway for protein degradation and is a pH dependent process. Peptides with glutamine as the N-terminal amino acid are known to undergo partial deamidation in solution, which converts the terminal glutamine residue into the cyclic pyroglutamyl residue (2-pyrrolidone-5-carboxylic acid).[2,3] Structural changes of the protein upon such reactions induce biological responses such as molecular clocks for the timing of biological processes or acceleration of amyloid aggregate formation (pathological protein formation).[4,5] Robinson et al.[5] strengthen the hypothesis that glutaminyl and asparaginyl residues in peptides serve, through deamidative transformation into glutamyl and aspartyl residues, as molecular timers of biological events such as protein development and aging. Amides may serve as molecular indicators that the integrity of each protein molecule has been maintained.[5] This fact can be used to investigate for the presence of post translational modifications.

A variety of studies have addressed the deamidation reactions of protonated Gln and Asn containing proteins and peptides in a mass spectrometer, where the reactions are initiated by collisional activation, occur in the gas phase, and can be analyzed in terms of the mass to charge ratio of the resulting fragments.[1,3,6-16] In order to gain detailed mechanistic insight into these reactions, the deamidation of protonated Asn and Gln has been studied using a variety of methods including energy-dependent collision-induced dissociation, deuterium exchange, and threshold collision induced dissociation (TCID) combined with theoretical investigations at various levels of theory.

While there is a general agreement that the thermodynamically most favored site of protonation of Gln and Asn is the α-amino group,[6,17] collisional activation can mobilize the additional proton,[18-20] so that elimination of NH_3 can conceivably occur from either the N-terminal amine or from the side-chain amide group. The latter option has been assumed in the majority of earlier studies[21] as it leads to an acylium ion structure – either as the final or intermediate product – rather than a presumably less stable α-carbocation. Although some studies suggest linear acylium-type product ions, most studies propose NH_3 loss to occur concomitantly with a nucleophilic attack onto the acylium or carbocation center by one of the nucleophiles in the molecule, leading to cyclic product ion structures. Schemes 3-1A and 3-1B outline the various possibilities for the formation of cyclic structures

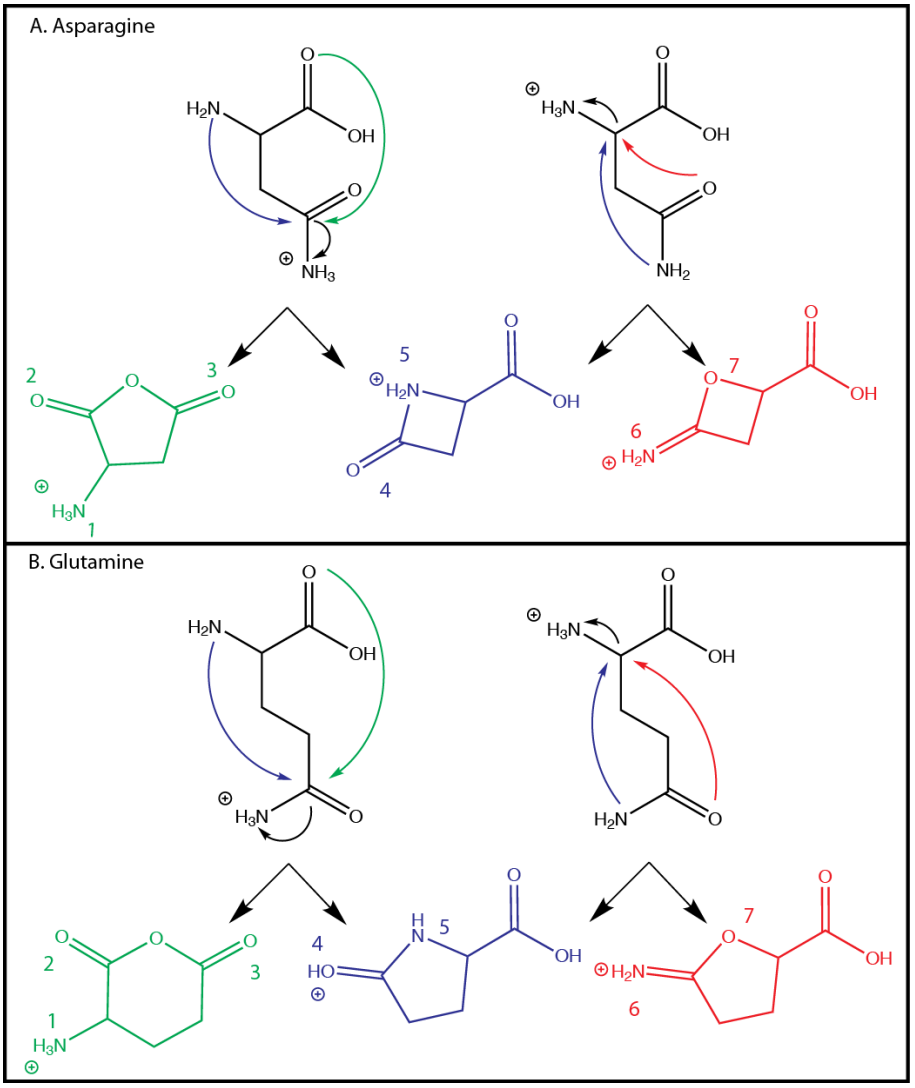
by deamidation of protonated Asn and Gln, respectively. For each of the conceivable product ion structures, several protonation sites are possible and further increase the diversity of the deamidation chemistry.

For $[Asn+H]^+$, Heaton and Armentrout[7] used TCID to obtain values for the activation energies of the various dissociation pathways, which were compared with detailed high-level computations of the potential energy surface. The study presented here confirms that NH_3 loss occurs from the amide side chain of the molecule as a result of nucleophilic attack by the C-terminal carbonyl oxygen and leads to a product ion having a succinic anhydride structure (see Scheme 3-1A). Other dissociation reactions, in particular the parallel reaction involving H_2O + CO loss from the C-terminus are also investigated here and will be addressed further below.

Number	$[Asn+H - NH_3]^+$ (kJ/mol)	$[Gln+H^+ - NH_3]^+$ (kJ/mol)
1	**0**	+ 61
2	+ 59	+ 100
3	+ 35	+ 92
4	+ 26	**0**
5	+ 32	+ 29
6	+ 64	+ 11
7	+ 188	+ 180

Table 3-1. *Calculated relative free energies at the MP2(full)/6-311+G(2d,2p)//B3LYP/6-31++G(d,p) level at 298 K for the different possible structures after NH_3 loss. The numbers refer to the different structures in Scheme 3-1.*

For $[Gln+H]^+$, the situation is less well understood; there is not an in depth study of the glutamine potential energy surface including transition-states and intermediate reaction species available. A reaction mechanism analogous to that identified for $[Asn+H]^+$ would lead to a six-membered glutaric anhydride structure. In contrast to $[Asn+H]^+$, however, the alternative mechanisms outlined in Scheme 3-1B lead to energetically more favorable five-membered ring structures. Table 3-1 outlines the free energy differences between these four- five and six-membered ring structures and their different protonation sites. Earlier studies[4,8-10,22] have indeed proposed such five-membered structures, in particular the oxo-proline structure (pyroglutamic acid). Further insights into the dissociation mechanisms may also be obtained from studies of the H_2O loss reaction of protonated glutamic acid $[Glu+H]^+$, which have often been suggested to lead to the same product ions.



Scheme 3-1:

A. *Conceivable reaction mechanisms for the loss of NH_3 from the side chain (blue and green) or from the N-terminus (red and blue) of $AsnH^+$. The numbers in the figure indicate the protonation site and correspond with the energies in table 3-1.*

B. *Conceivable reaction pathways for loss of NH_3 from $GlnH^+$. NH_3 can be expelled from the N-terminus or from the side-chain amide. Note that loss from the N-terminus by cyclization through the side-chain nitrogen and loss from the side chain by cyclization through the N-terminal nitrogen lead to the same oxo-proline product ion structure. The numbers in the figure indicate the protonation site and correspond with the numbers in Table 3-1.*

Recently it has been identified that residue specific fragmentation of protonated peptides containing glutamine or asparagine leading to succinimide and glutarimide structures, occurs[11,12] instead of the more standard formation of b_2 ions.[13] It was found that dissociation occurs via a nucleophilic attack of the side chain amide nitrogen atom of Asn and Gln on the adjacent peptide carbonyl carbon.

Ion spectroscopy is a proven tool for establishing the molecular structure of gas-phase amino acid ions[14], their (de)protonation site[23-25], and the structure of their gas-phase reaction products in CID MS/MS.[24] Here, the fragmentation reactions after CID of $[Asn+H]^+$ and $[Gln+H]^+$ are investigated using ion spectroscopy. Computed spectra for previously established or suggested isomers[7,8] are therefore compared with the measured IRMPD spectrum. Using the product ion structures we are able to propose fragmentation reaction mechanisms for deamidation of $[Asn+H]^+$ (confirming previously established pathways[7]) and $[Gln+H]^+$. These results confirm that $[Asn+H]^+$ and $[Gln+H]^+$ indeed follow different fragmentation pathways.

3.2. Experimental and computational methods

3.2.1. IRMPD spectroscopy

Infrared spectra of the protonated amino acids and their CID product ions were obtained by infrared multiple-photon dissociation (IRMPD)[26-28] spectroscopy using a Fourier Transform Ion Cyclotron Resonance mass spectrometer (FTICR-MS) as well as a modified 3D quadrupole ion trap MS (Bruker, AmaZon Speed ETD), both coupled to the beam line of the infrared free electron laser FELIX[27]. All samples were purchased from Sigma Aldrich (Zwijndrecht, The Netherlands) and used without further purification.

For experiments in the FTICR-MS, the protonated amino acids were generated by electrospray ionization (ESI) in a Micromass Z-spray source from approximately 1 mM solutions of the sample in a solution of 50:50 acetonitrile:water containing ~0.1% acetic acid. After the electrospray process, the ions were accumulated in a linear hexapole trap before being injected into the ICR cell. Fragment ions are generated by nozzle skimmer dissociation in the high-pressure region at the interface of the electrospray source and the hexapole trap/accumulation region. A stored waveform inverse Fourier transform (SWIFT)[29] excitation pulse is used to

isolate a single species in the cell. After irradiating the mass isolated ions for one or a few seconds with tunable infrared light (as described below), the mass spectrum of the ions which are left in the cell is recorded using an excite/detect procedure. The instrument and its typical operation has been described in detail elsewhere.[30]

For experiments in the Amazon ion trap MS, protonated amino acid ions were generated using ESI from 10^{-5} – 10^{-6} M solutions in 50:50 acetonitrile:water with ~0.1% formic acid. The yield at each IR frequency is obtained from three averaged mass spectra. Isolated ions were subjected to CID conditions for 40 ms using an amplitude parameter of approximately 0.3-0.5 V. Ions of interest were mass isolated and irradiated 1 – 4 pulses of tunable infrared light. A mass spectrum of the content of the trap is then obtained by scanning out the ion trap.

For both experiments, FELIX produced 6 μs long macropulses of approximately 20-60 mJ at a 5-Hz repetition rate having a bandwidth of ~0.5% of the center frequency. Absorption of infrared radiation increases the internal energy of the system which leads to frequency-dependent unimolecular dissociation. Relating parent and fragment ion intensities as the fragmentation yield (ΣI(fragment ions)/ΣI(parent + fragment ions)) and plotting the yield of each mass spectrum as a function of laser frequency generates an infrared vibrational spectrum. The yield is linearly corrected for the frequency dependent laser power and the infrared frequency is calibrated using a grating spectrometer.

3.2.2. Computational chemistry

Density Functional Theory (DFT) calculations were performed at the B3LYP/6-31++G(d,p) level of theory using Gaussian09 revision D01[31] to optimize the molecular geometries of the parent and fragment ions and to predict their (linear) IR spectra. Candidate input structures for these calculations were defined based on chemical intuition. The computed harmonic vibrational frequencies were scaled by 0.975 and convoluted with a 15 cm^{-1} full-width-at-half-maximum Gaussian line shape to facilitate comparison with experimental spectra.

A Molecular Mechanics/Molecular Dynamics (MM/MD) approach using AMBER 12 [32] was applied to protonated glutamine and glutamine-NH_3 in order to explore the potential energy surface and to find the lowest energy conformers. After minimization within AMBER, a simulated annealing procedure up to 1000 K was used. Guess geometries were first optimized in the Gaussian09 – revision D01 – package at the B3LYP/6-31++G(d,p) level of theory. 500 Structures were obtained

in this procedure and grouped based on structural similarity using appropriate rms criteria to give 20-30 candidate structures. Next, these structures were each optimized at the B3LYP/6-31++G(d,p) level as described above. The frequencies were again scaled by 0.975 and convoluted with a 15 cm^{-1} full-width-at-half-maximum Gaussian line shape. All reported 298 K free energies are calculated using MP2(full)/6-311+G(2d,2p) electronic energies with the B3LYP/6-31++G(d,p) optimized structures and thermochemistry. The computational procedure is described in more detail in ref.[15].

3.3. Results and discussion

3.3.1. Protonated asparagine $[Asn+H]^+$

The IRMPD spectrum of the $[Asn+H]^+$ precursor ion (m/z 133) has been recorded previously[16]. Comparison of the experimental spectrum with computed IR spectra for the three most stable conformers suggested that the amino acid is protonated at the N-terminus.

The fragmentation pathways of $[Asn+H]^+$ and $[Na+Asn]^+$, including the deamidation pathway, have also been investigated.[1,7] Two major routes for primary decomposition of $[Asn+H]^+$ were established: loss of NH_3 leading to an m/z 116 fragment and loss of H_2O + CO leading to an ion at m/z 87. A fragment at m/z 115 (loss of H_2O) was also found and suggested to be a result of nucleophilic attack of the side-chain oxygen atom onto the C-terminus. It was observed that the m/z 116 fragment dissociates further by loss of $O=C=CH_2$, leading to a $^+H_2N=CH-COOH$ fragment (m/z 74) and in parallel by loss of CO leading to $^+H_2N=CH-CH_2-COOH$ at m/z 88. The m/z 87 fragment can lose an HCN molecule giving $H_3N^+-C(=O)-CH_3$. Dookeran et al.[8] examined the fragmentation of $[Asn+H]^+$ and $[Gln+H]^+$ by energy-dependent CID using a ZAB-2FQ hybrid BEqQ mass spectrometer[33] by varying the collision gas. For $[Asn+H]^+$, low-energy metastable ion fragmentation was found to lead to dominant loss of H_2O and CO. At higher collision energies, loss of NH_3 and further fragmentation by elimination of ketene becomes dominant, leading to a linear $^+H_2N=CH-COOH$ fragment with m/z 74. Rogalewicz et al.[10] used a triple quadrupole mass spectrometer to examine the collision-induced dissociation of protonated α-amino acids with Argon at 12 eV lab energy (2.8 eV center of mass energy) under multiple collision conditions. From the estimated activation energy, the loss of NH_3 was predicted to occur from the side chain protonated amide intermediate, which is favored over NH_3 loss from the N-terminus since that involves formation of a destabilizing carbocation.[10]

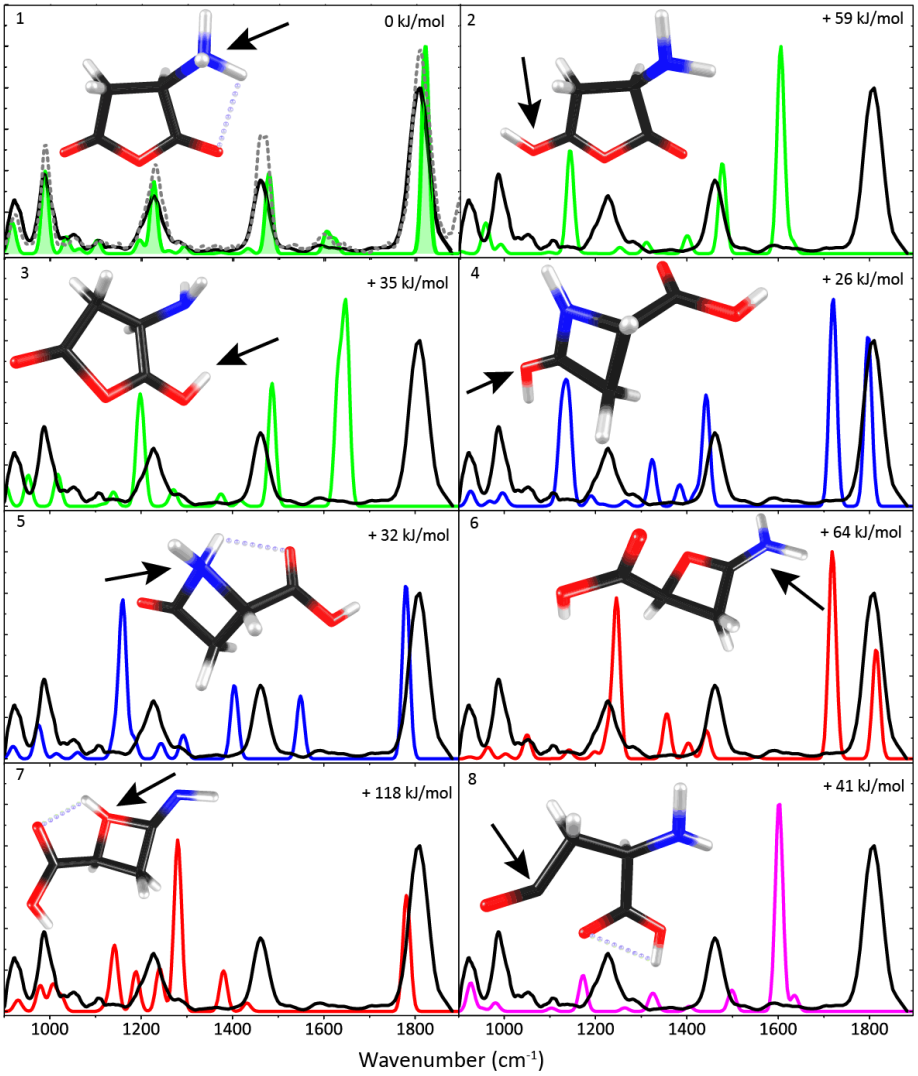


Figure 3-1. *Experimental IRMPD spectrum of the m/z 116 fragment ion from protonated Asn (in black) and DFT calculated spectra (colored) of the structures represented in Scheme 3-1. The IRMPD spectrum of the m/z 116 fragment ion from protonated Asp is shown as the grey dotted line in the top left panel and is seen to have an identical spectrum. The arrows in the figure indicate the protonation sites. Energy differences relative to the lowest energy structure are also shown. Structure* ***8*** *is the linear acylium ion, suggested in Refs.*[8,34]*. The shaded spectrum provides the best match with the experimental spectrum.*

Here, the structure of the fragments generated from CID of $[Asn+H]^+$ having m/z 116 and m/z 87 are examined using IRMPD ion spectroscopy. The intensity of the m/z 115 ion generated was too small for investigation. CID on the m/z 116 fragment resulted in m/z 88 m/z 74 and m/z 46 fragments.

Fragment: $[Asn+H-NH_3]^+$

Figure 3-1 shows the experimental IR spectrum of the mass-selected m/z 116 fragment ion along with the calculated spectra of the structures represented in Scheme 3-1A. The lowest energy structure found (see table 3-1), is the N-protonated 3-amino succinic anhydride structure (shown in the top left panel, and is indicated as (1) in Scheme 3-1A). This is the product ion structure identified in [7] based on threshold CID experiments and computation. The IR absorption band near 1820 cm^{-1} is assigned as the anti-symmetric combination of C=O stretches in the succinic anhydride structure. The band around 1625 cm^{-1} can be assigned to N-H bending, the band at 1480 cm^{-1} is assigned to C-H bending and the low-intensity bands at 1433 and 1374 cm^{-1} are attributed to CH_2 and C-H bending modes, respectively.

The four-membered ring protonated β-lactam carboxylic acid structure (structure **5** in Scheme 3-1A) was suggested to be the result of a nucleophilic attack by the terminal amino nitrogen on the side chain amide carbon atom.[10] The relative free energy of this isomer is +32 kJ/mol compared to the succinic anhydride isomer, although protonation at the lactam carbonyl oxygen (**4** in Scheme 3-1A) reduces the relative free energy difference to +26 kJ/mol. Nonetheless, the experimental IR bands around 1820 cm^{-1} and 1479 cm^{-1} are not well reproduced by the computed spectra for either of these lactam structures, as shown in Figure 3-1. The acylium ion structure reported in reference[8] is still higher in energy and can also be discarded based on the IR data.

Related: H_2O loss from protonated aspartic acid $[Asp+H]^+$

Loss of water is a well-known fragmentation channel for protonated amino acids with a carboxylic acid group in their side chain (Asp and Glu). Elimination of H_2O from protonated aspartic acid (m/z 134) leads to an ion at m/z 116 and the reaction mechanism has been suggested to be analogous to NH_3 loss from $[Asn+H]^+$, discussed above. Note that analogously to the suggested reaction pathways for protonated Asn, references[8,34] propose a linear acylium structure (see Figure 3-1) for the ion produced by loss of H_2O from protonated Asp.

Using IRMPD spectroscopy, it is now straightforward to verify whether the m/z 116 product ions from $[Asp+H]^+$ and $[Asn+H]^+$ possess the same molecular structure. The IR spectrum of $[Asp+H - H_2O]^+$, shown as dashed grey trace, is overlaid onto that of $[Asn+H\text{-}NH_3]^+$ in the top left panel of Figure 3-1. In terms of band frequencies, it is clear that all bands are reproduced within the experimental accuracy. We hence conclude that the loss of H_2O from protonated aspartic acid leads to an ion of the same structure as that of $[Asn+H\text{-}NH_3]^+$, i.e. a 3-amino-succinic anhydride structure. This does not correspond to the structure proposed in refs.[8,34].

Fragment: $[Asn+H–CO\text{-}H_2O]^+$

In addition to deamidation, loss of $CO+H_2O$ forms a major low-energy dissociation channel of $[Asn+H]^+$, leading to a product ion at m/z 87. This dissociation pathway dominates over NH_3 loss at the lowest collision energies.[8] Previous studies have explained this reaction pathway as being due to a proton transfer to the carboxyl group leading to sequential loss of H_2O and CO from the C-terminus, leading to a product ion with an immonium ion structure.[7] The full potential energy surface for the reaction was reported in ref.[7].

Figure 3-2 shows the experimental spectrum of the m/z 87 fragment ion from $[Asn+H]^+$ (black) along with the calculated spectra for the immonium ion structure (shaded purple) and alternative structures, the protonation sites are indicated by arrows. In the calculated spectrum of the immonium structure (2 in Figure 3-2), the band at 1700 cm^{-1} is assigned to the imine bond stretching mode, the peak at 1597 cm^{-1} is attributed to an amide NH_2 bending mode and the bands at 1371 cm^{-1} and 1123 cm^{-1} are CH_2 bending modes. The match with the computed IR spectrum for the immonium ion structure is rather good. Alternative structures are conceivable, *e.g.* different protonation sites or structures resulting from cyclization through nucleophilic attack by the side-chain nitrogen or oxygen atoms onto the β-carbon leading to four-membered ring structures. However, as shown in Figure 3-2, these alternatives are substantially higher in energy and feature computed IR spectra that do not match the experimental data.

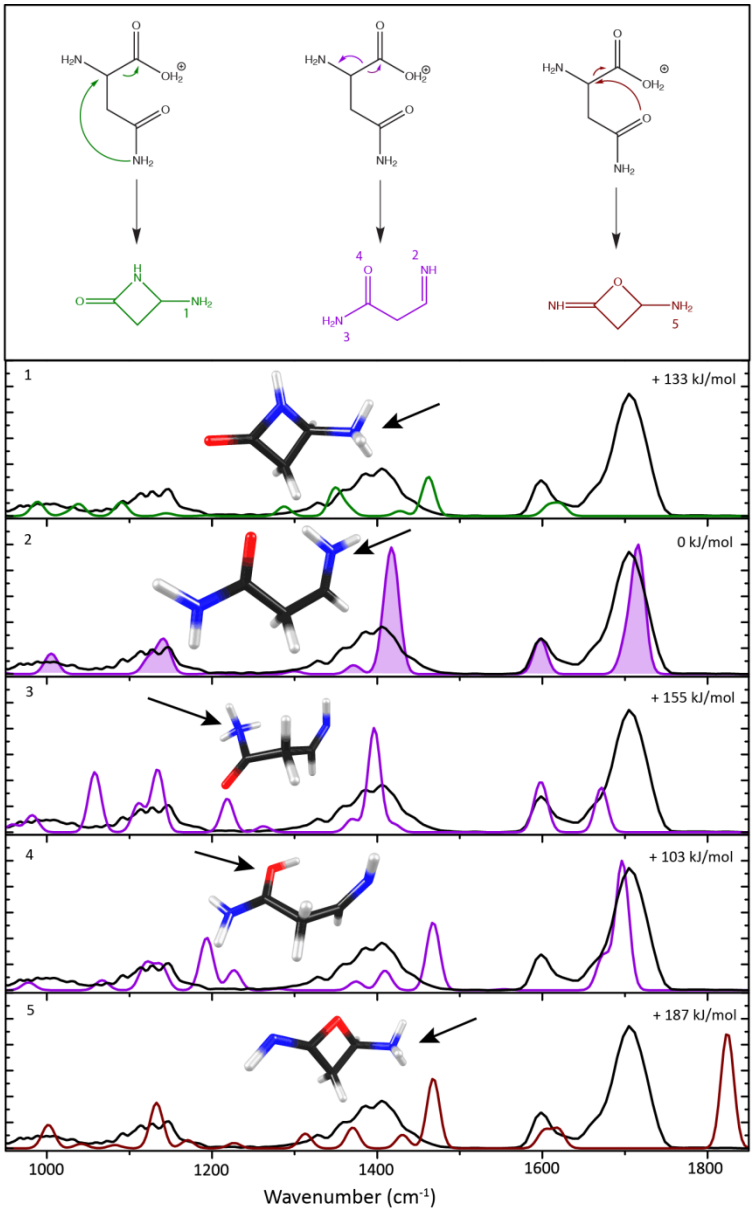
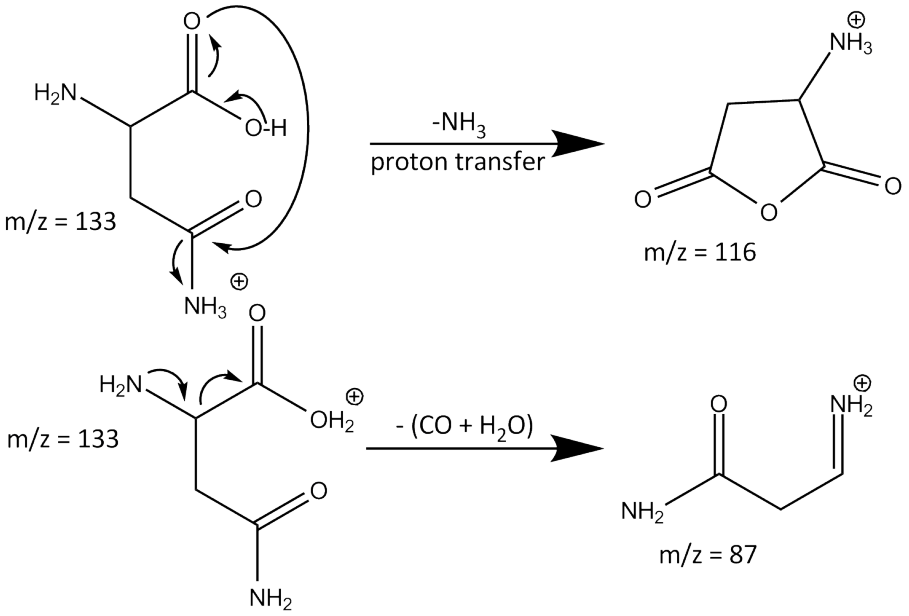


Figure 3-2. *Experimental IRMPD spectrum of the m/z 87 product ion of $[Asn+H]^+$ (in black) compared to the calculated DFT spectra of all reasonable structures for the m/z 87 fragment. The top panel presents all reasonable reaction mechanisms for the loss of a $CO+H_2O$ group (mass 46). $[Asn+H]^+$ can lose this group from the C-terminus after protonation of the carboxylic acid. The arrow indicates the protonation site and the shading indicates our assignment.*

Suggested reaction mechanism

The spectroscopically identified product ion structures suggest the reaction mechanisms shown in Scheme 3-2 for the loss of NH_3 and H_2O+CO. Loss of NH_3 occurs from the side chain amide by a nucleophilic attack of the C-terminal carbonyl oxygen, leading to a succinic anhydride structure. Loss of CO and H_2O occurs by detachment of the protonated C-terminal carboxylic acid group leaving behind an immonium ion, in accordance with what has been established previously[7].



Scheme 3-2. *Dissociation reaction mechanisms for $[Asn+H]^+$ as suggested from spectroscopically identified product ion structures.*

3.3.2. Protonated glutamine $[Gln+H]^+$

To verify the molecular structure of $[Gln+H]^+$ (m/z 147) in the gas phase, its IRMPD spectrum has been recorded (see Figure 3-3). The experimental spectrum is compared to computed spectra for two conformers of the N-terminally protonated molecule. Despite their non-negligible computed relative free energy difference of +17 kJ/mol, the experimental spectrum appears to match most closely to the calculated spectrum for the higher energy structure (blue shaded).

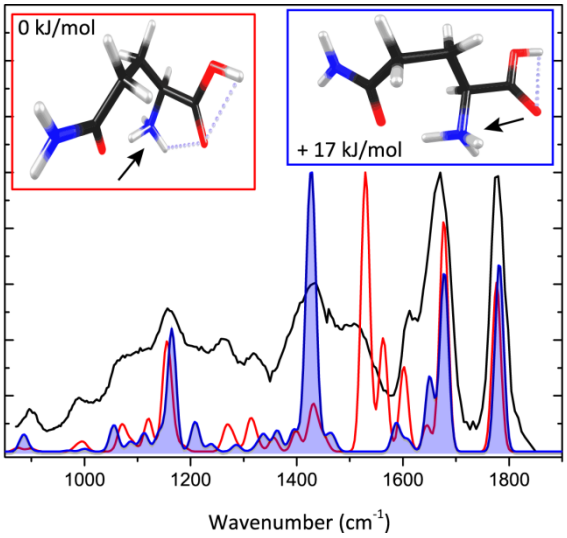


Figure 3-3. *Experimental IRMPD spectrum of $[Gln+H]^+$, m/z 147 (black) compared with computed spectra for the lowest energy conformer (red) and an alternative conformer at +17 kJ/mol (blue). The two conformers differ only slightly in the orientation and H-bonding of the protonated amino group. The higher energy conformer (blue shaded) appears to give the highest contribution to the experimental spectrum. This spectrum was measured in the FT-ICR MS. The arrows in the figure indicate the protonation site.*

The main structural difference between the two conformers lies in the orientation of the protonated amino terminus relative to the carboxyl and carbonyl oxygen atoms. In the lowest energy conformer, one of the amino protons H-bonds with the carbonyl oxygen and another one with the carboxyl oxygen. In the higher energy conformer, two protons H-bond with the carboxyl oxygen in a bifurcating fashion, while the third proton H-bonds with the carbonyl oxygen; the H-bonds in this conformer are significantly longer (1.6 and 2.4 Å) than in the lowest energy conformer (1.5 and 2.1 Å), indicating that the H-bonds are weaker and likely explaining the higher relative energy. The diagnostic absorption band at 1429 cm^{-1} in the higher energy conformer corresponds to the bending mode of the $–NH_3^+$ group. In the lowest energy conformer, the stronger H-bonding stiffens this vibrational mode, shifting it up in frequency to about 1530 cm^{-1}.

The two conformers have very similar structures and we expect that interconversion between them does not involve a significant barrier. It is therefore likely that a mixture of the two conformers exists and the spectrum appears to indicate this, particularly by the enhanced intensity in the 1400-1600 cm^{-1} range.

Spectral broadening of IRMPD bands is not uncommon in cases where structures can dynamically interconvert [23,35,36] and such a scenario is possibly at play here as well. The other features in the spectrum are assigned as the C=O stretch of the carboxylic acid group at 1781 cm^{-1}, the amide carbonyl stretch at 1677 cm^{-1}, modes having mainly N-H bending character at 1651, 1609, and 1587 cm^{-1}, and the O-H bending mode of the carboxylic acid group at 1164 cm^{-1}.

Fragment: $[Gln+H-NH_3]^+$

Various studies have reported that in contrast to $[Asn+H]^+$, elimination of NH_3 is the only dominant low-energy dissociation pathway for $[Gln+H]^+$, which is confirmed in our experiments. Elimination of NH_3 can conceivably occur from either the N-terminus or the side-chain amide group (see Scheme 3-1B). Detachment of the N-terminal amine would lead either to 2-imino-5-carboxy-tetrahydrofuran (blue structure in scheme 3-1B) if cyclization occurs through the side-chain oxygen or to 2-pyrrolidone-5-carboxylic acid (also referred to as pyroglutamic acid or oxo-proline) if cyclization occurs through the side-chain nitrogen (red structure in scheme 3-1B). Elimination of NH_3 from the side-chain amide has been considered to be more likely, and will lead to the same oxo-proline structure (blue in scheme 3-1B) if cyclization occurs through the N-terminal nitrogen. If cyclization occurs through the C-terminal carboxyl oxygen, a 3-amino-glutaric anhydride (green structure in scheme 1B) is formed. The latter reaction is analogous to that identified for NH_3 loss from $[Asn+H]^+$, leading instead to the 5-membered succinic anhydride structure. The energies of these suggested structures are listed in Table 3-1.

Figure 3-4 shows the experimental IRMPD spectrum for the m/z 130 fragment corresponding to $[Gln+H-NH_3]^+$ (black trace). The blue trace represents the vibrational spectrum predicted for the oxo-proline structure proposed in earlier studies [8,37] and identified as the deamidation product of sodiated glutamine [22]. Computed spectra for protonation on the lactam nitrogen (4 in scheme 3-1B) atom as well as on the oxo O-atom (5 in scheme 3-1B) are shown. In the calculated spectrum, the band at 1786 cm^{-1} is assigned to C=O stretching of the carboxylic acid group, the absorption at 1687 cm^{-1} can be attributed to lactam CN stretching. The bands at 1480, 1419, 1276, 1198 and 1140 cm^{-1} have predominantly C-H bending character. Changing the protonation site to the nitrogen atom (Figure 3-4B) yields a predicted spectrum that does not match the experiment, as is most obvious from the bands at 1786 and 1687 cm^{-1}. The relative free energy is +29 kJ/mol (see table 3-1).

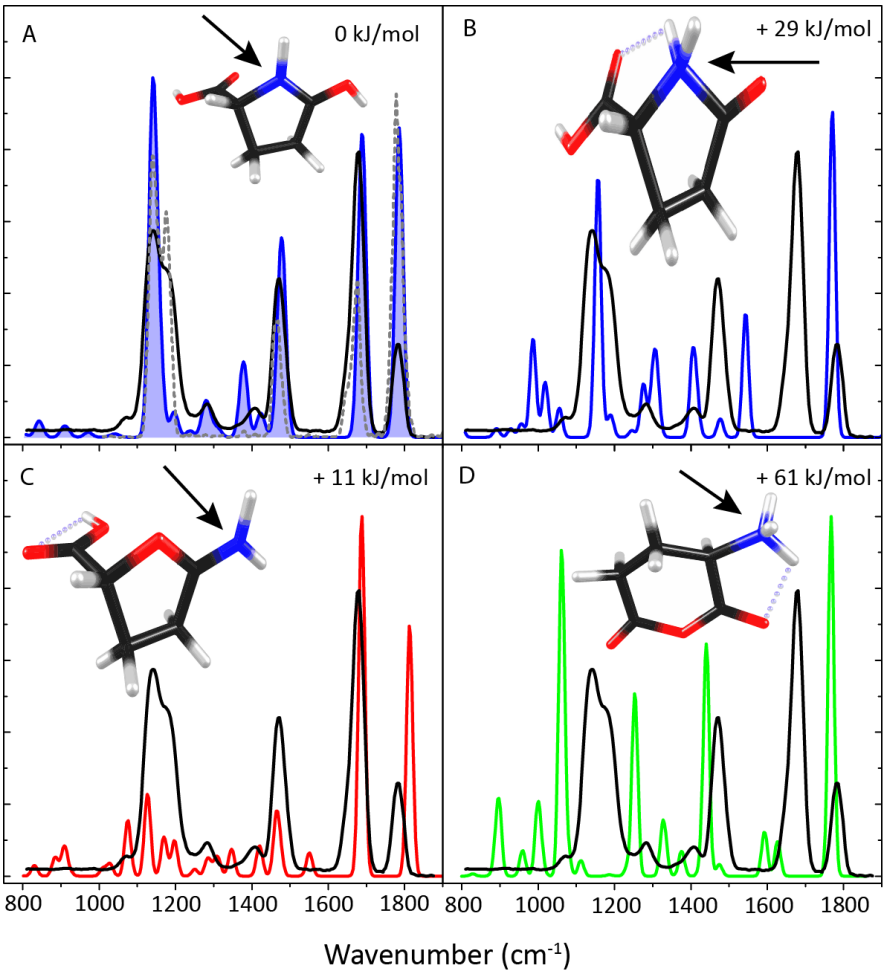


Figure 3-4. *Experimental IRMPD spectrum of $[Gln+H-NH_3]^+$ (black in all panels) compared with calculated spectra (colored according to the structures in Scheme 3-1B) for different isomeric structures for the m/z 130 ion. The best match is identified for the oxo-O protonated oxo-proline structure (top left, is the assigned calculation and therefore shaded). The dotted grey line represents the experimental spectrum for the m/z 130 fragment from $[Glu+H]^+$ (H_2O loss). The arrows indicate the protonation site of the structure.*

To firmly establish the oxo-proline structure for $[Gln+H - NH_3]^+$ (4 in scheme 3-1B) alternative structures were considered by comparing their predicted spectra to the experimental one. If the NH_3-loss mechanisms for $[Asn+H]^+$ and $[Gln+H]^+$ were analogous, a six-membered amino-glutaric anhydride structure (structure 1 Scheme 3-1B) would have resulted for deamidation of $[Gln+H]^+$. The relative free energy of this structure is 60-100 kJ/mol higher in energy (dependent on the

protonation site, see table 3-1) than the oxo-proline structure. Calculated spectra for this structure do not match the experimentally observed IR spectrum, for instance, the experimental bands at 1786 and 1687 cm^{-1} are not reproduced in the calculated spectra (see Figure 3-4D). A final alternative is the substituted tetrahydrofuran structure resulting from elimination of the N-terminal amine and concomitant cyclization through the side-chain oxygen atom (Figure 3-4C, 6 in scheme 3-1B). The relative free energy of this structure is only +11 kJ/mol higher than the oxo-proline structure. The calculated IR spectrum for this species, however, does not match as well as that for the O-protonated oxo-proline structure, especially considering the band at 1786 cm^{-1} and around 1100-1200 cm^{-1}. From this we assign the the oxo-O-protonated oxo-proline structure as the product ion from deamidation of $[Gln+H]^+$.

Related: H_2O loss from protonated glutamic acid $[Glu+H]^+$

Protonated glutamic acid ($[Glu+H]^+$, m/z 148) is similar to $[Gln+H]^+$ except for a carboxylic acid group in the side chain replacing the amide group. Elimination of H_2O from $[Glu+H]^+$ has been suggested [8,37] to lead to the same product ion as that formed by NH_3 loss from $[Gln+H]^+$. In order to verify this, the IRMPD spectrum of $[Glu+H-H_2O]^+$ (grey trace) is overlaid onto the IRMPD spectrum of $[Gln+H-NH_3]^+$ (black trace) in Figure 3-4. Although the relative intensities of the bands show some deviations, attributed to differences in the experimental conditions, the band frequencies (1786, 1681, 1480 and 1140 cm^{-1}) are reproduced closely, leading us to conclude that the product ions from $[Glu+H-H_2O]^+$ and $[Gln+H-NH_3]^+$ are the same.

Fragment: $[Gln+H - NH_3 - H_2O - CO]$

In contrast to $[Asn+H]^+$, loss of H_2O + CO from $[Gln+H]^+$ is not observed in parallel with deamidation. Instead, the $[Gln+H-NH_3]^+$ fragment ion undergoes subsequent loss of H_2O + CO to arrive at a product ion at m/z 84. In [8] it is proposed that CO and H_2O are formed from elimination of the carboxylic acid group from protonated oxo-proline leading to a pyrolidone-like structure shown in Figure 3-5. Based on deuterium labeling experiments, an alternative acyclic structure for the m/z 84 fragment ion from $[Gln+H]^+$ and $[Glu+H]^+$ has also been suggested [37]. Computed spectra for both isomers are shown in Figure 3-5 along with the experimental IR spectrum for the m/z 84 fragment from $[Gln+H-NH_3]^+$. We suspect that the small size of the molecule in combination with its rigidity causes the multiple-photon absorption and dissociation process to be relatively inefficient, so that weaker bands (especially at 1250 and 1050 cm^{-1}) are not observed in the experimental IRMPD spectrum.

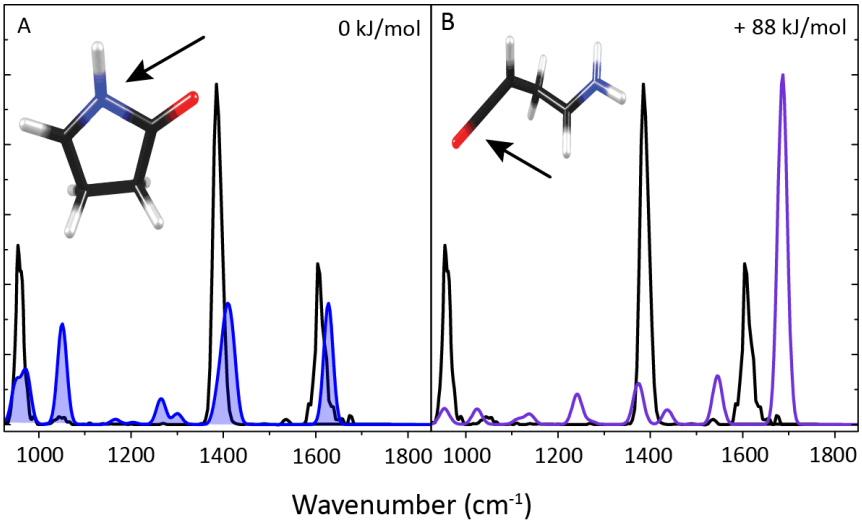


Figure 3-5. *Experimental IRMPD spectrum of the m/z 84 fragment from $GlnH^+$ - NH_3 (black) compared with calculated spectra for cyclic (A) (blue) and acyclic (B) (purple) isomers of this fragment. The cyclic pyrolidone-like structure is seen to provide the closest match to the experimental spectrum. The arrow indicates the protonation sites of the cyclic pyrolidone-like structure, for the acyclic structure, the arrow indicates the charged site.*

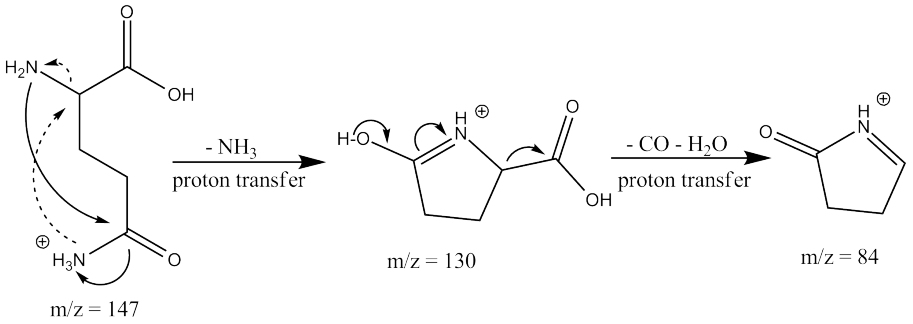
The m/z 84 fragment is a result of the detachment of CO and H_2O after loss of an NH_3 from $[Gln+H]^+$. This is in contrast to the fragmentation mechanism of protonated asparagine, where the m/z 87 ion is formed directly from asparagine. Attributing the experimental structure to the cyclic isomer in Figure 3-5A, the peak at 1627 cm^{-1} is assigned to the C-N stretching in the lactam. The peak around 1410 cm^{-1} is assigned to C-H bending, the peaks at 1050 and 973 cm^{-1} to C-C stretching. The only other plausible alternative is the same pyrolidone-like structure with protonation at the oxygen instead of at the nitrogen, which gives a structure that is 57 kJ/mol higher in energy. The computed spectrum for this protomer clearly disagrees with the experimental spectrum (see Supporting Information Figure S3 in ref. [38]), so that we assign the structure in Figure 3-5A to the m/z 84 fragment ion.

Suggested reaction mechanism

From the identified structures in Figures 3-4 and 3-5, a suggested reaction mechanism can be derived, as presented in Scheme 3-3. Proton transfer from the N-terminus to the side-chain amide nitrogen induces elimination of a neutral ammonia molecule with a concomitant nucleophilic attack of the N-terminal nitrogen onto the carbonyl carbon of the side chain. This leads to an oxo-proline structure protonated on the oxo-oxygen atom (shown in its resonant form in

Scheme 3-3). Further activation transfers the proton to the hydroxyl group of the carboxylic acid eliminating H_2O + CO. This leaves behind a fragment ion at m/z 84 with a pyrolidone-like structure.

Admittedly, the identification of the structure of the m/z 130 ion does not distinguish between a reaction mechanism in which the ammonia molecule is eliminated from the N-terminus or from the side-chain amide. Both reaction pathways lead to the same oxo-proline structure. Various reports have suggested NH_3 to be eliminated from the side chain as it leads to an intermediate carbonium ion (at the side-chain carbonyl) rather than an intermediate carbocation (at the α-carbon), which would occur upon NH_3 loss from the N-terminus and which is estimated to be less stable.



Scheme 3-3. *The reaction mechanism (solid arrow) for the fragmentation of $[Gln+H]^+$ as suggested from the product ion structures identified from their IRMPD spectra. The dashed arrow indicates an alternative mechanism, starting from the structure with the proton on the N-terminus, but leading to the same product ion structure.*

3.4. Conclusion

We have presented infrared spectra for CID products of $[Asn+H]^+$ and $[Gln+H]^+$, including a comprehensive characterization of the product ions from the deamidation reactions. With the assistance of quantum-chemical computations, these spectra identify the product ion structures. $[Asn+H-NH_3]^+$ is thus shown to possess a 3-amino succinic anhydride structure protonated on the amino group, as had been identified on the basis of measured and computed activation energies. Likewise, the parallel dissociation reaction involving simultaneous loss of H_2O and CO is confirmed to form a linear immonium ion.

The infrared ion spectrum for the NH_3-loss product from $[Gln+H]^+$ establishes that a different reaction mechanism occurs for this species. While neutral ammonia is again eliminated from the side chain amide, the N-terminal nitrogen now acts as the nucleophile leading to an oxo-proline product ion structure. Both $[Asn+H]^+$ and $[Gln+H]^+$ thus form 5-membered ring structures upon deamidation. The fragmentation reactions for H_2O loss from protonated aspartic and glutamic acid have also been spectroscopically investigated. These fragmentation reactions lead to structures that are identical to those found for the loss of NH_3 from $[Asn+H]^+$ and $[Gln+H]^+$.

Chapter 4

Deamidation reactions of asparagine and glutamine containing dipeptides investigated by ion spectroscopy

Deamidation is a major fragmentation channel upon activation by collision induced dissociation (CID) for protonated peptides containing glutamine (Gln) and asparagine (Asn) residues. Here, we investigate these NH_3-loss reactions for four Asn- and Gln-containing protonated peptides in terms of the resulting product ion structures using infrared ion spectroscopy with the free electron laser FELIX. The influence of the side chain length (Asn vs. Gln) and of the amino acid sequence on the deamidation reaction has been examined. Molecular structures for the product ions are determined by comparison of experimental IR spectra with spectra predicted by density functional theory (DFT). The reaction mechanisms identified for the four dipeptides AlaAsn, AsnAla, AlaGln and GlnAla are not the same. For all four dipeptides, primary deamidation takes place from the amide side chain (and not from the N-terminus) and in most cases resembles the mechanisms previously identified for the protonated amino acids asparagine and glutamine. Secondary fragmentation reactions of the deamidation products have also been characterized and provide further insight in – and confirmation of – the identified mechanisms. Overall, this study provides a comprehensive molecular structure map of the deamidation chemistry of this series of dipeptides[*].

[*] Adapted from: Kempkes, L. J. M.; Martens, J.; Grzetic, J.; Berden, G.; Oomens, J., Deamidation Reactions of Asparagine- and Glutamine-Containing Dipeptides Investigated by Ion Spectroscopy. *J. Am. Soc. Mass Spectrom.* **2016,** *27* (11), 1855-1869.

4.1. Introduction

Under natural conditions, deamidation of glutamine (Gln) and asparagine (Asn) residues in proteins is a primary route for post-translational modifications, playing a role in several diseases, and is believed to relate to aging effects acting as a molecular clock.[4,5,39-41] Studies of the mechanistics of Asn and Gln deamidation in condensed media have shown that the reaction proceeds via cyclic imide intermediates, *i.e.* glutarimide for Gln and succinimide for Asn.[39-41] These imides are formed via nucleophilic attack by the adjacent peptidyl amide nitrogen onto the side chain carbonyl carbon atom with concomitant elimination of NH_3 (green arrows in Scheme 4-1). Alternatively, for peptides with Gln as the N-terminal residue, partial deamidation in solution occurs by conversion of the Gln residue into a cyclic pyroglutamyl residue (2-pyrrolidinone-5-carboxylic acid).[2,3,42] This type of reaction is indicated for GlnAla (QA) with the light and dark brown arrows in Scheme 4-1.

Also in the gas phase, *i.e.* as protonated ions in a mass spectrometer, Gln and Asn containing peptides are subject to facile NH_3 loss under activation by collision induced dissociation (CID). In fact, loss of small neutral molecules including in particular NH_3 and H_2O commonly competes with backbone cleavage under slow heating conditions employed in MS-based peptide sequencing.[43] To better understand the mechanistics involved in deamidation of protonated peptides containing Gln and Asn, we employ here infrared ion spectroscopy to resolve the structures of the product ions resulting from NH_3-loss from the protonated dipeptides of AlaAsn, AsnAla, AlaGln and GlnAla. The sensitivity and extensive MS^n capabilities of the ion trap mass spectrometer used in this study enable not only the characterization of MS/MS reaction products, but also of products resulting from further stages of CID (MS^n). We thus construct a deep molecular structure map of the reaction network, where structural identification of product ions at different MS^n stages provides further consistency for the entire network.

In recent years, ion spectroscopy has emerged as a useful probe of ion structures of small gas-phase species and their reaction products after CID, being capable of distinguishing between isomers, protomers and even conformers.[24,25,30,38,44-53] It has for instance been recently shown that b_2-sequence ions from specific peptides containing Gln or Asn do not form along the common oxazolone dissociation pathway, nor via the alternative diketopiperazine pathway.[11] Although yet other alternatives had been suggested earlier[39,54-56], cyclic imide structures formed

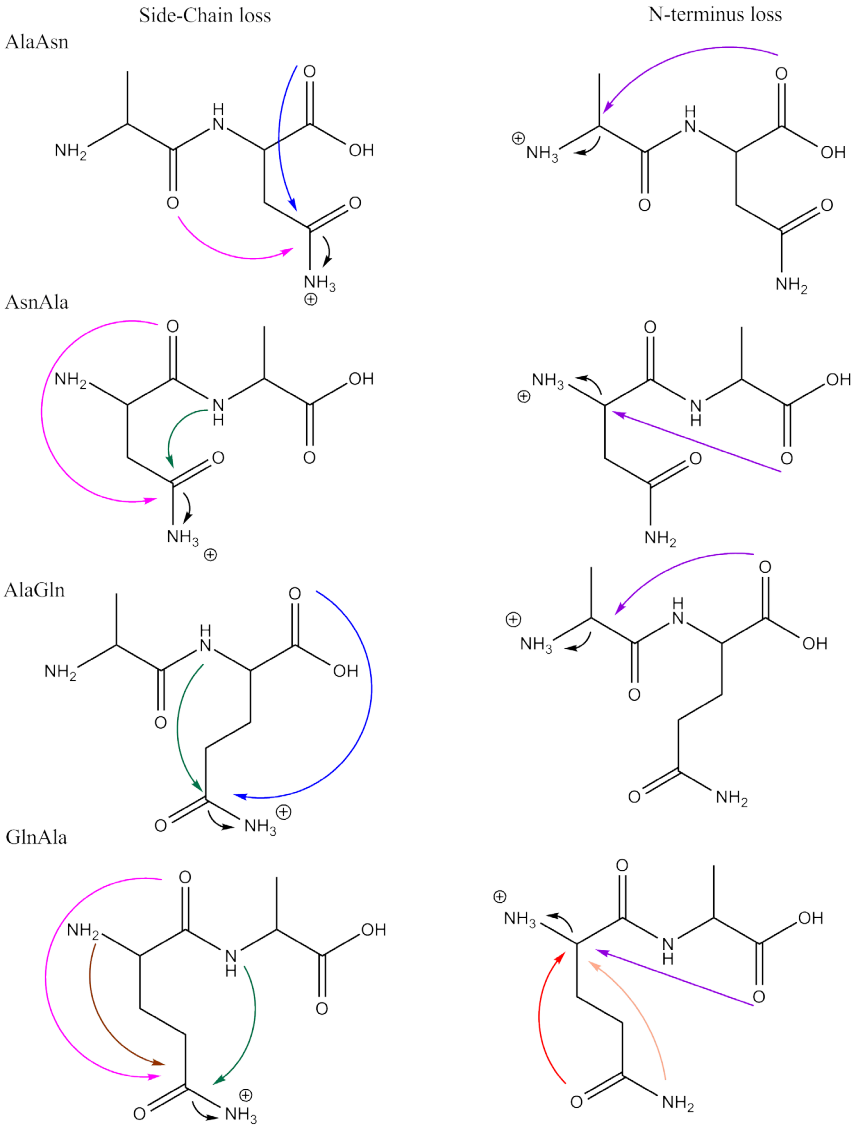
through nucleophilic attack by the amide nitrogen atom of the Gln and Asn side chains was firmly established only after analysis of IR spectra of the product ions.

The mobile proton model for peptide dissociation proposes that the added proton, initially being localized on the most basic site, migrates upon mild collisional activation to a less favored protonation site.[6,18-20,57] For this reason, the deamidation reaction for the investigated dipeptides can conceptually take place from either the N-terminus or the side-chain amide. For the protonated amino acids of Gln and Asn, it was recently established that NH_3-loss occurs in both cases from the amide side chain and that deamidation is assisted by a nucleophilic attack leading to cyclization (see chapter 3 and ref.[7]). Nonetheless, the reaction mechanism is not the same for the two amino acids, as the attacking nucleophile is the C-terminal carbonyl O-atom for Asn and the N-terminal N-atom for Gln. Following these reaction pathways, a 5-membered ring product ion is formed for both amino acids: 3-amino-succinic anhydride is formed from protonated Asn and pyroglutamic acid (2-pyrrolidone-5-carboxylic acid) is formed from protonated Gln.

Scheme 4-1 outlines for each of the four dipeptides studied here the possibilities for the formation of 5- and 6-membered ring structures after NH_3-loss from either the amide side chain or the N-terminus. Attacks leading to 4-, 7-, or 8-membered ring structures were found to be energetically disfavored and to provide no agreement with the experimental IR spectra; in the interest of conciseness, we will not consider those structures further in the text of this report, although they were considered during the analysis of the spectra. As can be concluded from Scheme 4-1, for each of the dipeptides, three or more isomeric $[M + H - NH_3]^+$ structures are conceivable and moreover, for each of them multiple protomers and conformers are generally possible, illustrating in general the complexity of the problem.

For protonated AlaAsn and AlaGln, the blue arrow in Scheme 4-1 indicates a deamidation mechanism analogous to that established for the protonated Asn amino acid (see chapter 3 and ref.[7]). Similarly for protonated GlnAla, nucleophilic attack of the N-terminal nitrogen onto the side chain amide carbon (dark brown arrow) is analogous to the mechanism observed for the Gln amino acid in chapter 3. Note that nucleophilic attack by the side-chain nitrogen onto the α-carbon of the N-terminal residue (light brown arrow) leads to the same product and would thus be indistinguishable by spectroscopic characterization of the fragment ion; a similar ambiguity was encountered for the Gln amino acid. For AsnAla, some of the mechanisms analogous to those indicated for GlnAla would result in 4-membered ring species and are not included in Scheme 4-1.

Notable mechanistic studies, both experimental and theoretical, on the deamidation reactions occurring in the gas phase upon CID of protonated peptides include those reported in Refs.[3,9,42,58,59]. We note that these reactions have often been considered to be analogous to dehydration reactions from Glu and Asp containing peptides.[9,34] A relatively large number of studies have addressed NH_3-loss from peptides with an N-terminal Gln residue[2,3,9,42], where the consensus is that the deamidation of the Gln side chain leads to the formation of a pyroglutamyl moiety following the brown arrow in the bottom left panel of Scheme 4-1. Deamidation of protonated AlaGln is discussed by Harrison[9] and considered analogous to dehydration of AlaGlu, although the structure of the resulting fragments could not be determined with certainty. The fragmentation chemistry of protonated ValAsn and ValGln have been investigated by theoretical modeling at the density functional theory level.[58] Deamidation reaction mechanisms were predicted to be analogous for the two peptides, occurring from the side chain by a nucleophilic attack of the backbone amide oxygen onto the side chain carbonyl carbon. In Scheme 4-1, this mechanism is indicated for AlaAsn by the pink arrow; the reactions lead to 6- and 7-membered lactone ring structures for ValAsn and ValGln, respectively. Our analysis of deamidation of protonated AlaGln and AlaAsn presented below suggest a different mechanism. Armentrout and coworkers very recently reported a detailed investigation of the potential energy surface for NH_3-loss from protonated AsnGly[59], establishing reaction pathways similar to those found here for protonated AsnAla.



Scheme 4-1. *Possible reaction mechanisms for NH_3 loss from the protonated dipeptides AlaAsn, AsnAla, AlaGln and GlnAla. Both ammonia loss from the side chain (left column) as well as from the N-terminus (right column) are taken into account. This Scheme only considers NH_3 loss induced by nucleophilic attack leading to 5- or 6-membered ring products. Mechanisms leading to energetically less favorable 4-,7- or higher membered ring structures have been considered but are not shown here. Analogous reactions are indicated by arrows of the same color. The dark and light brown arrows for GlnAla indicate two different reactions leading to the same product.*

4.2. Experimental and computational methods

4.2.1. IRMPD spectroscopy

The experiments were performed in a modified 3D quadrupole ion trap mass spectrometer (Bruker, AmaZon Speed ETD) coupled to one of the beam ports of the Free Electron Laser for Infrared eXperiments (FELIX).[15,27,60] Protonated peptides are generated by electrospray ionization (ESI) and fragments are generated by collision induced dissociation (CID) and mass-selected using the standard MS^n routines of the instrument.

Peptide samples were purchased from GeneCust (Luxemburg) and used without further purification. For each experiment, a freshly prepared solution was used to avoid naturally occurring degradation of the dipeptides in the solution. Protonated dipeptide ions were generated using ESI from 10^{-5} – 10^{-6} M solutions in 50:50 acetonitrile:water and ~0.1% formic acid. Deamidated ions were obtained by subjecting the precursor ions to CID conditions for 40 ms using an amplitude parameter of approximately 0.3 – 0.5 V. Sequential fragmentation is induced by subjecting the mass-isolated deamidated $[M + H - NH_3]^+$ ions to a second stage of CID (MS^n).

To record infrared spectra of the deamidated dipeptides and further fragmented product ions, InfraRed Multiple Photon Dissocation (IRMPD) spectroscopy[26,28] has been applied after mass-isolation of the product ion of interest using the IR radiation from FELIX. Infrared absorption by the system causes an increase in the internal energy, which leads to frequency-dependent uni-molecular dissociation. The dissociation yield at each IR wavelength is determined from three averaged mass spectra. An infrared vibrational spectrum is generated by relating the precursor and IR induced fragment ion intensities according to

$$Yield = \sum I_{fragment} \Big/ (\sum I_{fragment} + I_{precursor})$$

where the sum is over all fragments produced by IR irradiation. An assumed linear correction is applied for the frequency dependent variations in the laser power. Tunable infrared radiation from FELIX is delivered in pulses of approximately 6 µs duration, 20 – 60 mJ energy and at a bandwidth of ~0.5% of the center frequency at a repetition rate of 10 Hz. The frequency is calibrated using a grating spectrometer.

The IRMPD spectrum of $[AlaGln+H-NH_3]^+$ is measured using the combination of FELIX with a Fourier Transform Ion Cyclotron Resonance mass spectrometer (FTICR-MS).[61,62] The experimental procedure is similar to that described above and has been detailed elsewhere.[30,50]

4.2.2. Computational chemistry

IR spectra are predicted and the molecular geometries are optimized for isomeric structures that were deemed possible by chemical intuition, including all likely protomers. Calculations on these first guess structures were performed at the Density Functional Theory (DFT) level with the B3LYP/6-31++G(d,p) basis set, using Gaussion09 revision D01.[31]

For the lowest energy structures of these $[M + H - NH_3]^+$ fragments, potential energy surfaces are further explored to identify the lowest energy conformers using a Molecular Mechanics/Molecular Dynamics (MM/MD) approach employing AMBER 12.[32] After minimization within AMBER, a simulated annealing procedure up to 500 K was used. Five hundred structures were obtained in this procedure and grouped based on structural similarity using appropriate rms criteria to give 20-30 candidate structures. Next, these structures were each optimized at the B3LYP/6-31++G(d,p) level as described above. A more detailed description of the procedure can be found elsewhere.[15] For all optimizations on all geometries, the calculations were completed with electronic energy calculations at the MP2(full)/6-311+G(2d,2p)//B3LYP/6-31++G(d,p) level at 298 K.

For the purpose of comparison with experimental spectra, the DFT computed harmonic vibrational frequencies were scaled by 0.975 and convoluted with a 20 cm^{-1} full-width-at-half-maximum Gaussian line shape.

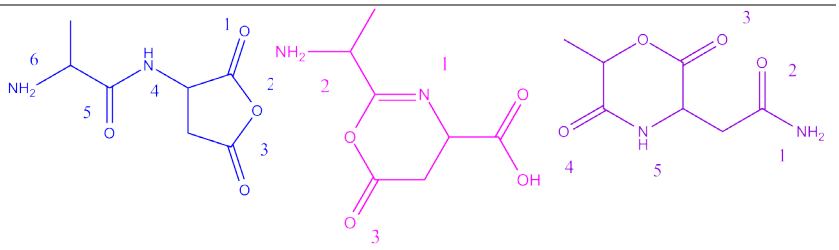
4.3. Results and discussion

4.3.1. Deamidated Alanine-Asparagine $[AlaAsn+H-NH_3]^+$

As depicted in Scheme 4-1, three nucleophilic attacks are considered most likely to occur in the deamidation reaction of $[AlaAsn+H]^+$. The C-terminal carbonyl oxygen can attack the asparagine side chain, leading to the 5-membered succinic anhydride structure **1** in Table 4-1. Alternatively, the oxygen from the amide linkage can attack the asparagine side chain, leading to a 6-membered ring structure, **2**, which has been proposed on the basis of a computational study for the deamidation of protonated ValAsn.[58] A third possibility involves loss of NH_3

from the N-terminus, resulting from an attack by the C-terminal carbonyl oxygen on the N-terminus, leading to the morpholine ring structure **3**.

For each of the resulting structures, multiple protonation sites are conceivable. The most likely ones are indicated with numbers in Table 4-1 and their relative computed energies are given. The 5-membered succinic anhydride structure, **1.6**, protonated at the N-terminus is the overall lowest energy structure.



Name → H^+ site ↓	1 succinic anhydride	2 dihydro-oxazinone	3 diketo-morpholine
1	+ 160	+ 8	+ 66
2	+ 180	+ 17	+ 47
3	+ 176	+ 131	+ 68
4	+ 124		+ 50
5	+ 36		+ 54
6	**0**		

Table 4-1: Deamidation of protonated AlaAsn. *Calculated relative free energies at the MP2(full)/6-311+G(2d,2p)//B3LYP/6-31++G(d,p) level at 298 K for the different possible isomeric structures after NH_3 loss, for each of which different protonation sites have been considered. The colors of the structures correspond with the colors of the arrows in Scheme 4-1, for example, the nucleophilic attack of the C-terminal oxygen on the side chain is indicated with a blue arrow, leading to structure **1** presented in blue in this table.*

The left panel of Figure 4-1 shows the experimental IRMPD spectrum (black) of the m/z 187 deamidation product ion $[AlaAsn+H-NH_3]^+$ compared with the computed spectra of **1.6** in two nearly iso-energetic conformations (grey and blue). Figure S1 in ref.[63] shows the experimental spectrum compared with computed spectra for each of the suggested structures in Table 4-1. Based on spectral matching, we propose to assign the anhydride structure (**1.6**) in blue as the structure of the ion present in the experiment (spectra of assigned structures are shaded throughout the paper). Note that this structural assignment is at odds with the dissociation mechanism proposed for ValAsn[58], which would lead to the oxazinone structure **2.1**; clearly, the computed spectrum for this structure does not match the experimental one (see Figure S1 in ref.[63]). The difference between the spectra of the two conformers of **1.6** is most prominent for the band near 1530 cm^{-1}, which is due to the N-H bending normal mode of the amide bond. For the identified structure shown in blue, the bands around 1820 and 1900 cm^{-1} are assigned as the symmetric and anti-symmetric combination of C=O stretches of the succinic anhydride ring; the band around 1700 cm^{-1} is due to C=O stretching of the amide carbonyl.

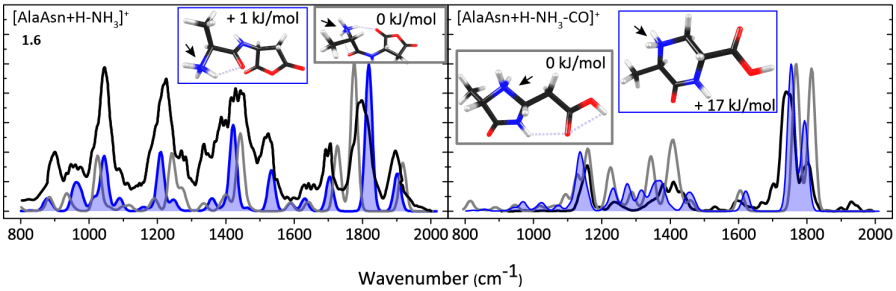
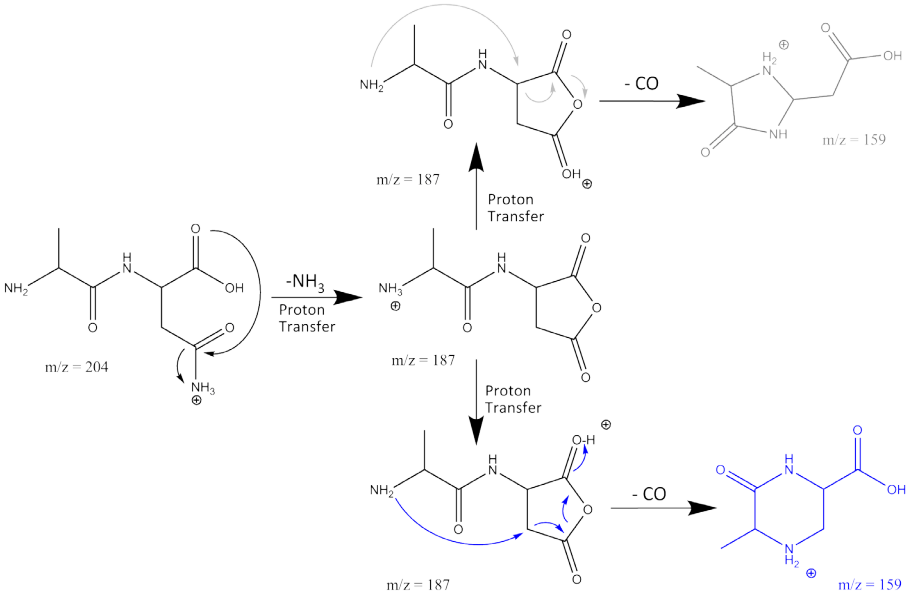


Figure 4-1. *The experimental IRMPD spectra (black) of AN-17 (left panel) and AN-17-28 (right panel) compared to the best matching (blue) and the lowest energy (grey) calculated spectra. Of the two nearly iso-energetic **1.6** conformers in the left panel, we suspect the structure in blue to have the highest contribution to the experimental IRMPD spectrum (indicated by the shading). The protonation sites are indicated with an arrow.*

Isolation and an additional stage of CID on the $[AlaAsn+H-NH_3]^+$ ion generates mainly a fragment at m/z 159, corresponding to $[AlaAsn+H-NH_3-CO]^+$. We use IRMPD spectroscopy on this MS^3 fragment ion in order to establish its structure, which can also provide further confirmation of the structure identified for $[AlaAsn+H-NH_3]^+$ above. The loss of 28 mass units can correspond to neutral loss of CO or C_2H_4. While the resolution of the mass spectrometer is not sufficient to distinguish between them, inspection of the proposed structure **1.6** for [AlaAsn+H-

$NH_3]^+$ suggests that loss of CO is more plausible than loss of C_2H_4. The right panel of Figure 4-1 shows the experimental IRMPD spectrum of $[AlaAsn+H-NH_3-CO]^+$ together with calculated spectra for two low energy isomers (grey and blue) resulting from CO loss from **1.6**. The blue-framed structure is the product of a nucleophilic attack by the N-terminal nitrogen onto the CH_2-group of the succinic anhydride ring structure, leading to loss of CO from the ring and rearrangement to a 6-membered ring structure. The grey-framed structure is the result of a nucleophilic attack by the N-terminal nitrogen onto the CH in the succinic anhydride ring, leading to a 5-membered ring structure.



Scheme 4-2. *The suggested reaction mechanism for the fragmentation of protonated AlaAsn as suggested from the product ion structures identified from their IRMPD spectra.*

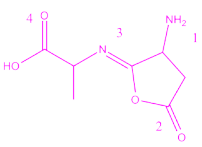
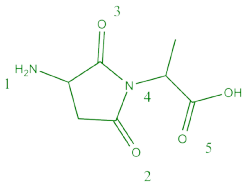
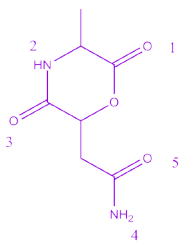
For the $[AlaAsn+H-NH_3-CO]^+$ spectrum in grey, the vibration calculated at 1810 cm^{-1} is due to stretching of the carbonyl group in the ring, while that at 1760 cm^{-1} is due to the carboxyl C=O stretch. The weak band around 1600 cm^{-1} is due to NH_2 bending. For the alternative $[AlaAsn+H-NH_3-CO]^+$ structure (blue), the band near 1800 cm^{-1} is due to the carboxyl C=O stretch, while that at 1760 cm^{-1} is identified as C=O stretching of the carbonyl in the 6-membered ring. The peak around 1620 cm^{-1} is due to NH_2-bending of the protonated secondary nitrogen. While the experimental spectrum appears to match slightly better with the computed spectrum for the higher energy 6-membered ring structure, the similarity of the

two calculated spectra does not truly allow us to identify one or the other and the spectrum may also be due to a mixture of both structures. Nonetheless, both isomers can conceivably be formed from the $[AlaAsn+H-NH_3]^+$ structure identified above, so that the analysis of the IR spectrum for $[AlaAsn+H-NH_3-CO]^+$ provides additional support for the structural assignment of **1.6** as the $[AlaAsn+H-NH_3]^+$ fragment. Scheme 4-2 shows the reaction mechanisms for the formation of $[AlaAsn+H-NH_3]^+$ and $[AlaAsn+H-NH_3-CO]^+$, including pathways leading to both suggested isomers for the secondary neutral loss of 28 mass units.

4.3.2. Deamidated Asparagine-Alanine $[AsnAla+H-NH_3]^+$

Scheme 4-1 presents three possible reaction mechanisms for the deamidation of $[AsnAla+H]^+$. The amide linkage oxygen can attack the side chain leading to a 5-membered lactone ring, **4** in Table 4-2, the nitrogen in the amide linkage can attack the side chain leading to a succinimide ring, **5,** or the C-terminal oxygen can attack the N-terminal α-carbon leading to **6**. For each isomer, multiple protonation sites are again conceivable and the most prominent ones are shown with their relative computed energies in Table 4-2. Structure **5.1** represents the lowest-energy structure.

Upon IRMPD of $[AsnAla+H-NH_3]^+$, fragments are produced at m/z 170 $[AsnAla+H-NH_3-NH_3]^+$ and at m/z 169 $[AsnAla+H-NH_3-H_2O]^+$ and it was noticed that the spectral responses in each of these channels are significantly different. This is suggestive of there being two different isomers of the precursor $[AsnAla+H-NH_3]^+$ ion, each having a different IR spectrum and undergoing different fragmentation (i.e. neutral loss of either NH_3 or H_2O). The top left panel of Figure 4-2 shows the experimental spectrum of $[AsnAla+H-NH_3]^+$ as derived from the mass channel corresponding to sequential loss of 17, the bottom left panel shows the IRMPD spectrum derived from sequential loss of 18. The two spectra are analyzed by comparison with computed spectra for the series of isomers and protomers listed in Table 4-2. The best match for the spectrum observed in mass channel 170 is provided by the computed spectrum of lactone structure **4.3** (pink), while the spectrum observed in mass channel 169 is best reproduced by the computed spectrum for succinimide structure **5.1** (green). These structures match with those recently established for deamidation of protonated AsnGly on the basis of TCID measurements and extensive theoretical modeling [59]. No match was found with spectra computed for the morpholine ring structures **6**, as further shown for its lowest energy protomer **6.3** in Figure S2 in ref. [63].

Name → **H$^+$ site ↓**	**4** **imino-butyrolactone**	**5** **succinimide**	**6** **diketo-morpholine**
1	+ 110	**0**	+ 81
2	+ 138	+ 36	+ 152
3	**+ 46**	+ 82	+ 25
4	+ 189	+ 154	+ 86
5		+ 104	+ 46

Table 4-2. Deamidation of protonated AsnAla. *Calculated relative free energies at 298 K for different isomeric structures and protonation sites after loss of NH_3. Colors of the suggested structures correspond to colors of the arrows showing the reaction mechanism in Scheme 4-1.*

The most striking difference between the two individual experimental spectra for protonated $[AsnAla+H-NH_3]^+$ is the band at 1955 cm^{-1}, which is assigned to the stretching of the carbonyl group in the lactone ring structure of **4.3**. More subtle differences are observed for the doublet of bands in the 1600 – 1800 cm^{-1} range; the computed spectra for the lactone **4.3** and succinimide **5.1** structures closely reproduce these differences. The most prominent deviation between experiment and theory is the ~50 cm^{-1} shift encountered for the band near 1500 cm^{-1} in the spectrum of the succinimide structure (green), which is due to the NH_3 bending mode. The computed spectrum of **4.3** (pink) indicates that the band around 1775 cm^{-1} is due to C=O stretching of the carboxyl group, the band at 1690 cm^{-1} corresponds to C=N stretching of the imine bond and the weaker band near 1630 cm^{-1} is due to NH_2 bending of the amine. For **5.1**, the computations indicate that the band around 1766 cm^{-1} is the anti-symmetric combination of succinimide C=O stretches, the band at 1695 cm^{-1} is due to the C=O stretch of the carboxyl group

and the bands at 1644 cm^{-1}, 1604 cm^{-1} and 1507 cm^{-1} are due to bending modes of the NH_3 group.

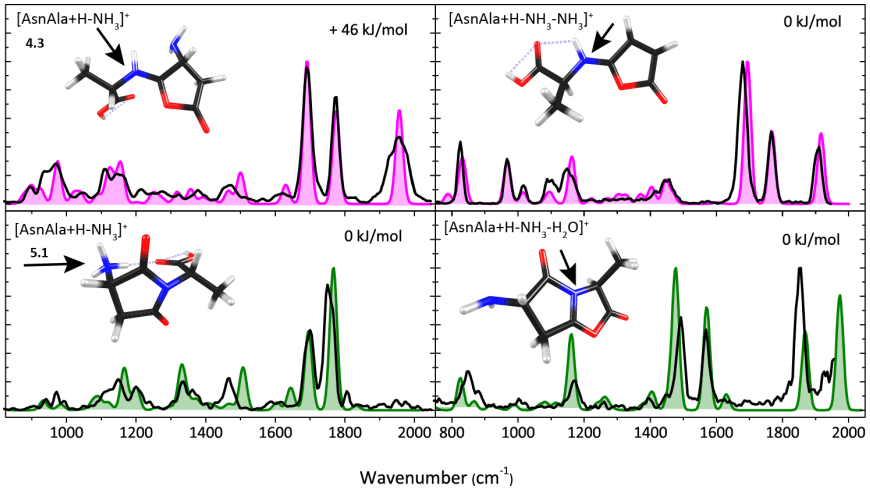
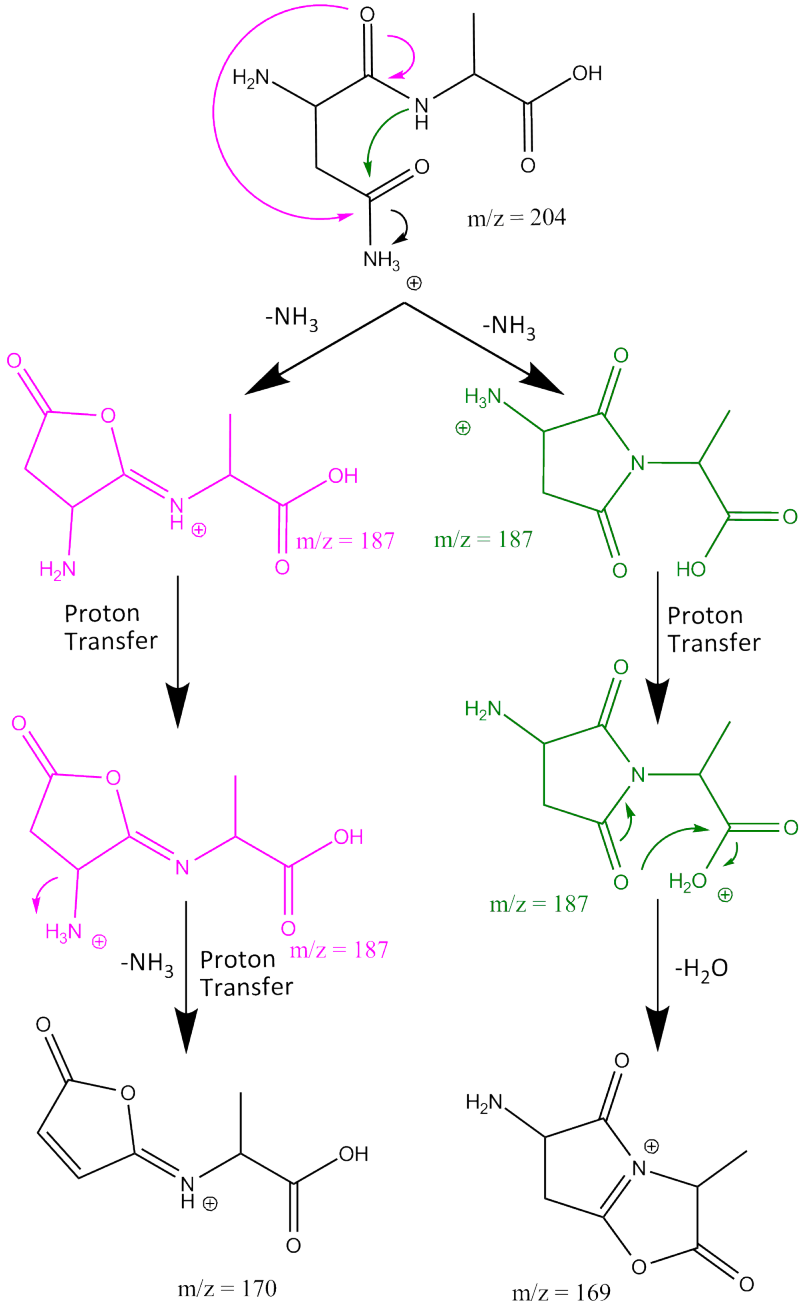


Figure 4-2. *Experimental IRMPD spectra of* $[AsnAla+H-NH_3]^+$ *(left panel) ,* $[AsnAla+H-NH_3-NH_3]^+$ *(top right panel) and* $[AsnAla+H-NH_3-H_2O]^+$ *(bottom right panel) compared to the best matching calculations. For* $[AsnAla+H-NH_3]^+$*, two isomers are identified (top panel and lower panel) by separate analysis of the IR induced dissociation from* $[AsnAla+H-NH_3]^+$ *into mass channels m/z 169 (H_2O loss) and m/z 170 (NH_3 loss) and are assigned to structures* ***4.3*** *(calculated spectrum in pink) and* ***5.1*** *(green). IRMPD spectra of the two product ions* $[AsnAla+H-NH_3-NH_3]^+$ *and* $[AsnAla+H-NH_3-H_2O]^+$ *have also been recorded (black traces in the panels on the right). The protonation sites are indicated with an arrow.*

IRMPD spectroscopy was then applied to the MS^3 fragments $[AsnAla+H-NH_3-NH_3]^+$ and $[AsnAla+H-NH_3-H_2O]^+$ to further follow the dissociation cascade and to support the structural assignment for the two isomers of $[AsnAla+H-NH_3]^+$ above. The spectrum of the $[AsnAla+H-NH_3-NH_3]^+$ (m/z 170) MS^3 fragment is shown in the top right panel of Figure 4-2 and matches excellently with the spectrum predicted for an imine substituted furanone structure, with protonation on the imine nitrogen (see Scheme 4-3). This suggests that **4.3** undergoes further fragmentation by loss of another ammonia molecule by direct cleavage of the amine substituent on the lactone ring. In contrast, structure **5.1** undergoes subsequent loss of H_2O. The IRMPD spectrum of the $[AsnAla+H-NH_3-H_2O]^+$ (m/z 169) fragment is well reproduced by the spectrum predicted for a double 5-membered ring structure shown in the lower right panel of Figure 4-2.



Scheme 4-3. *Suggested reaction mechanism for the fragmentation of protonated AsnAla as derived from the product ion structures identified from their IRMPD spectra. For* [AsnAla+H-NH_3]$^+$*, two isomers are found, each exhibiting different secondary breakdown reactions.*

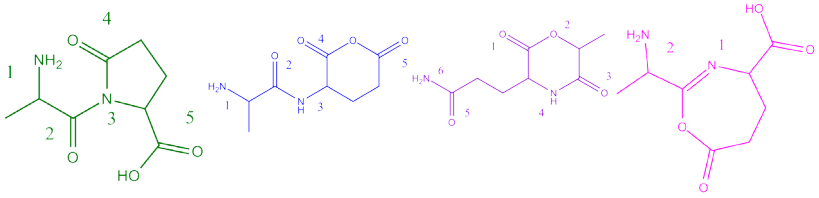
The identification of this fragment structure suggests that water is detached from the carboxylic acid group with concomitant nucleophilic attack of one of the succinimide carbonyl oxygen atoms onto the carboxylic carbon (see Scheme 4-3). Recording the spectra of the MS^2 and MS^3 fragments thus yields a complete and consistent picture of the deamidation reactions, as summarized in Scheme 4-3.

For $[AsnAla+H-NH_3-NH_3]^+$, the band around 1915 cm^{-1} is assigned to C=O stretching in the carbonyl group and the band around 1470 cm^{-1} is assigned to C-H bending in the methyl group of the molecule. For the calculated $[AsnAla+H-NH_3-H_2O]^+$ spectrum, the peak around 1973 cm^{-1} is assigned to C=O stretching in the oxazolone ring, while the peak at 1868 cm^{-1} is assigned to the C=O stretch of the pyrolidone ring. The peak at 1629 cm^{-1} is assigned to NH_2 bending, the peak at 1570 cm^{-1} is assigned to C=N stretching, the peak at 1477 cm^{-1} is assigned to C-O stretching in the lactone ring, and the peak at 1160 cm^{-1} is due to C-H bending.

The identification of the various structures leads to a suggested reaction mechanism as shown in Scheme 4-3. For the pathway depicted in pink, the oxygen at the amide linkage attacks the asparagine side chain inducing the detachment of the first ammonia molecule. This leads to the imino-butyrolactone structure **4** analogous to the deamidation reaction of the amino acid asparagine (chapter 3). The second ammonia molecule is lost by cleavage of the amine substituent on the lactone ring after transfer of the proton from the imine nitrogen. For the parallel reaction depicted in green in Scheme 4-3, the first step is a nucleophilic attack of the amide bond nitrogen on the asparagine side chain, leading to the succinimide structure **5**, which was also observed in the deamidation reaction of asparagine containing peptides in solution.[39-41] Subsequent fragmentation involves the loss of a water molecule from the C-terminus.

4.3.3. Deamidated Alanine-Glutamine $[AlaGln+H-NH_3]^+$

Three reasonable reaction pathways for the deamidation of protonated AlaGln are shown in Scheme 4-1. The C-terminal oxygen can attack the glutamine side chain leading to **8,** characterized by a glutaric anhydride ring; the amide bond nitrogen can attack the glutamine side chain resulting in **7**, having a 5-membered γ-lactam ring; and NH_3 can be lost from the N-terminus by concomitant nucleophilic attack by the C-terminal oxygen leading to the diketo-morpholine structure **9**. Although structure **10** has a 7-membered ring, we have included it here as it was previously suggested as the NH_3-loss product of protonated ValGln.[58]



Name → H^+ site ↓	7 γ-Lactam	8 Glutaric anhydride	9 Diketo-morpholine	10
1	**0**	+ 64	+ 140	+ 126
2	+ 83	+ 85	+ 223	+ 132
3	+ 123	+ 177	+ 90	
4	+ 98	+ 154	+ 102	
5	+ 132	+ 167	+ 130	
6			+ 184	

Table 4-3: Deamidation of protonated AlaGln. *Calculated relative free energies at 298 K for different possible isomers generated by NH_3 loss, each having various possible protonation sites.*

Table 4-3 shows that the 5-membered lactam ring structure with the added proton at the N-terminus (**7.1**) has the lowest energy. Alternative structures are in this case substantially higher in energy, with the glutaric anhydride structure protonated at the N-terminus (**8.1**) being the next higher isomer at +64 kJ/mol. Figure 4-3 shows the calculated absorption spectra of structures **7.1** (two nearly iso-energetic conformers in grey and green) and **8.1** (blue) compared to the experimental IRMPD spectrum (in black).

Structure **8.1** would be formed if the deamidation of $[AlaGln+H]^+$ followed a reaction mechanism analogous to that identified above for deamidation of $[AlaAsn+H]^+$. However, as observed clearly in the bottom left panel of Figure 4-3, the predicted IR spectrum for **8.1** does not match with the experimental spectrum. We conclude that $[AlaGln+H]^+$ detaches NH_3 by a different mechanism as compared to its asparagine analogue. Furthermore, no match was found with computed

spectra for the various protomers of **9** nor with those for the previously suggested structure **10**, featuring a 7-membered ring. Figure S3 in ref.[63] shows spectral comparisons with the computed spectra for **9.3, 10.1** and **10.2**.

The top left panel of Figure 4-3 shows the experimental spectrum of [AlaGln+H-NH_3]$^+$ compared with computed spectra for two conformers of the lowest energy protomer of the lactam ring containing product (**7.1**), having computed relative free energies within 2 kJ/mol. For the conformer shown in green, characterized by H-bonding between the protonated N-terminus and the amide carbonyl (rather than with the lactam carbonyl for the conformer shown in gray), the match between experimental and computed spectra is reasonable. The 1780 cm^{-1} band is assigned as being due to the two unresolved C=O bands in the lactam ring and in the carboxyl group, computed at 1781 and 1783 cm^{-1}. The peak at 1680 cm^{-1} is due to the N-terminal ammonium group. We cannot exclude the presence of a minor fraction of the alternative conformer (grey) in the ion population, especially given the absorption observed around 1720 and 1460 cm^{-1}. Further evidence for the assignment of the γ-lactam ring structure is provided by analysis of the structure of the MS^3 fragment ion.

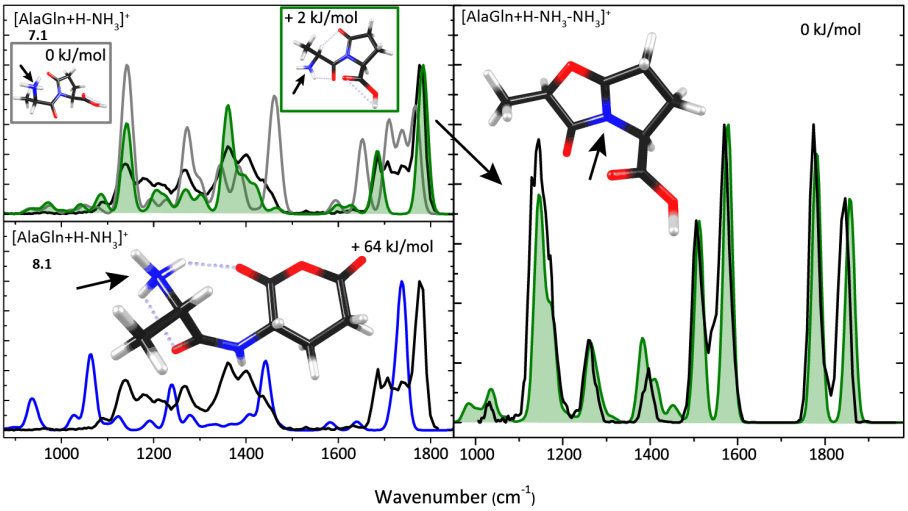
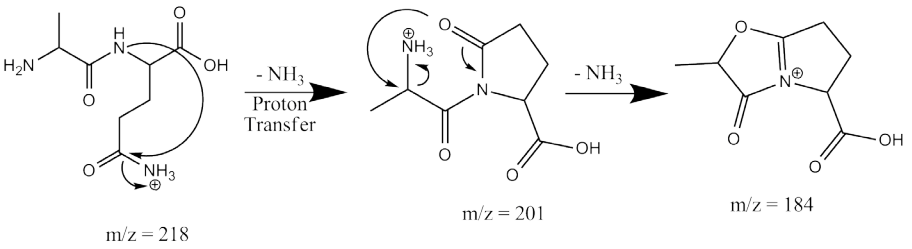


Figure 4-3. *Experimental IRMPD spectrum of* [AlaGln+H-NH_3]$^+$ *(left panels, black) compared with computed spectra for structures* ***7.1*** *(two conformers in green and grey) and* ***8.1*** *(blue). The right panel shows the experimental IRMPD spectrum of* [AlaGln+H-NH_3-NH_3]$^+$ *(black) along with the spectrum predicted for a structure having expelled* NH_3 *from the N-terminus of* ***7.1*** *with concomitant ring closure by nucleophilic attack of the succinimide oxygen onto the* α*-carbon of the Ala residue. Protonation sites are indicated with an arrow.*

Activation of the isolated m/z 201 fragment induces the loss of a second NH_3 molecule giving m/z 184 $[AlaGln+H-NH_3-NH_3]^+$, in line with observed MS^3 behavior in Ref.[9]. Detachment of the N-terminal ammonium group from **7.1** with concomitant nucleophilic attack of the lactam oxygen onto the α-carbon of the Ala residue leads to a bicyclic structure shown in the right panel of Figure 4-3. The calculated spectrum of this $[AlaGln+H-NH_3-NH_3]^+$ fragment is shown in green and matches closely with the experimental spectrum. The bands at 1860 and 1780 cm^{-1} are due to the C=O stretch modes of the oxazolone and carboxyl moieties, respectively, while the band at 1580 cm^{-1} is due to stretching of the C=N bond joining the two 5-membered rings. The bands around 1400-1500 cm^{-1} are due to several C-H bending modes in the molecule.

Identification of the bicyclic structure as the CID product of ammonia loss from $[AlaGln+H-NH_3]^+$ supports our above identification of the precursor $[AlaGln+H-NH_3]^+$ species as the lactam-containing ion **7.1.** Scheme 4-4 summarizes the proposed reaction mechanism, which is based on the spectroscopically identified structures. This sequence of side-chain and N-terminal ammonia loss is in agreement with what has been suggested by Harrison[9], although the structure of the MS^3 fragment could not be determined in that study.



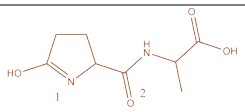
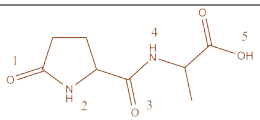
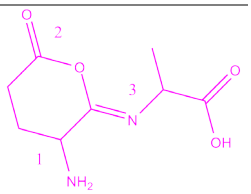
Scheme 4-4. *Reaction mechanism for the fragmentation of $[AlaGln+H]^+$ as suggested from the product ion structures identified from their IRMPD spectra.*

4.3.4 Deamidated Glutamine-Alanine $[GlnAla+H-NH_3]^+$

Of the dipeptides studied here, GlnAla has the largest number of possible deamidation reaction mechanisms leading to 5- or 6-membered ring structures. Table 4-4 lists the 6 structures resulting from the suggested reaction mechanisms in the bottom row of Scheme 4-1. Nucleophilic attack by the N-terminal nitrogen onto the side chain and nucleophilic attack of the side chain nitrogen onto the α-carbon of the Gln residue lead to isomerically identical structures, although they can exist

in two tautomeric forms indicated as **11** (enol) and **12** (keto) in Table 4-4. In its keto form, this structure is usually referred to as the pyroglutamyl product often suggested to result from NH_3-loss from peptides with an N-terminal Gln residue [2,3,9,42]. Depending on their site of protonation, the two tautomers may become equivalent, in particular, structures **11.1** and **12.1** are equivalent. The 6-membered δ-valerolactone derivative **13** is the result of a nucleophilic attack by the oxygen of the amide linkage onto the side chain. A nucleophilic attack of the amide bond nitrogen on the side chain leads to the glutarimide structure **14**; this mechanism has been suggested for dehydration of peptides containing an N-terminal Glu residue [9]. Loss of NH_3 can also occur from the N-terminus through attack of the side-chain oxygen leading to **15** or through attack of the C-terminal oxygen leading to **16**.

The left panel in Figure 4-4 shows the experimental spectrum for the $[GlnAla+H-NH_3]^+$ product ion along with computed spectra for two nearly iso-energetic conformers of the lowest energy isomer **11.1** in Table 4-4 (see Fig. S4 in ref.[63] for a comparison with other structures listed in Table 4-4). The 0 kJ/mol conformation has a slightly more compact structure than the conformer computed at +2 kJ/mol. The compact conformation is induced by an H-bond between the carboxylic carbonyl oxygen and the NH of the pyroglutamyl ring, while in the more extended conformer this O-atom forms an H-bond with the amide NH. The overall shape of the spectroscopic feature observed between 1600 and 1800 cm^{-1} suggests that the extended conformer, whose predicted spectrum is shown in brown in the left panel of Figure 4-4, is the dominant contributor to the observed spectrum. On the other hand, the compact structure, whose predicted spectrum is shown in gray, cannot be excluded and especially the feature observed near 1200 cm^{-1} suggests a significant contribution. For the extended structure of $[GlnAla+H-NH_3]^+$, the band at 1760 cm^{-1} is due to the carboxyl C=O stretch, the band at 1710 cm^{-1} to C=O stretching of the peptide bond and the band at 1680 cm^{-1} is due to C=N stretching in the pyroglutamyl moiety. The bands around 1500 cm^{-1} are attributed to C-H bending.

			
Name →	**11**	**12**	**13**
H^+ site ↓	**Tautomerized pyroglutamyl**	**Pyroglutamyl**	**Imino-valerolactone**
1	**0**	**0**	+ 95
2	+ 91	+ 42	+ 161
3		+ 54	+ 124
4		+ 111	
5		+ 74	

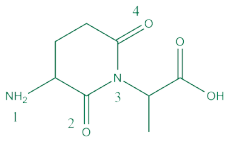
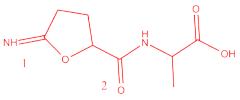
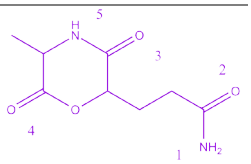
			
Name →	**14**	**15**	**16**
H^+ site ↓	**Glutarimide**	**Imino-dihydrofuran**	**Diketo-morpholine**
1	+ 27	+ 59	+ 78
2	+ 33	+ 114	+ 24
3	+ 126		+ 93
4	+ 51		+ 103
5			+ 125

Table 4-4: Deamidation of [GlnAla+H]$^+$. *Calculated relative free energies at 298 K for different isomeric structures possibly resulting from NH_3-loss, for different protonation sites for each of them. Note that **11** and **12** are tautomers of the same isomer; depending on the protonation site, the two structures may be equivalent, i.e. **11.1** is the same as **12.1**.*

MS^3 of $[GlnAla+H-NH_3]^+$ did not provide $[GlnAla+H-NH_3-NH_3]^+$ as observed for $[AlaGln+H-NH_3]^+$. This is in line with the $[GlnAla-NH_3]^+$ structure identified on the basis of the spectrum in the left panel of Figure 4-4, from which it is not obvious how to expel another NH_3. Instead, the main reaction channel observed involves H_2O loss giving $[GlnAla+H-NH_3-H_2O]^+$. The structure of this MS^3 fragment is also determined by IRMPD spectroscopy as shown in the right panel of Figure 4-4. Water loss from the C-terminus is conceivable after H^+-migration and nucleophilic attack by the amide carbonyl oxygen onto the C-terminal carboxyl carbon. This leads to the formation of an oxazolone ring protonated on the oxazolone nitrogen. Alternative protonation sites produce structures with relative free energies that are at least 20 kJ/mol higher. For this γ-lactam-oxazolone bicyclic structure of $[GlnAla+H-NH_3-H_2O]^+$, a spectrum is predicted with strong bands at 1950 cm^{-1} due to the oxazolone C=O stretch, at 1800 cm^{-1} due to the lactam carbonyl stretch, at 1610 cm^{-1} due to oxazolone CN stretching and at 1450 cm^{-1} due to C-H bending in the molecule. This predicted spectrum matches the experimental spectrum for $[GlnAla+H-NH_3-H_2O]^+$ reasonably well so that we can summarize the breakdown of $[GlnAla+H]^+$ as shown in Scheme 4-5. Note that for the NH_3-loss in the first reaction step leading to the well-known pyroglutamyl ring, a nucleophilic attack from the N-terminus on the side-chain is spectroscopically indistinguishable from a nucleophilic attack from the side chain to the N-terminus.

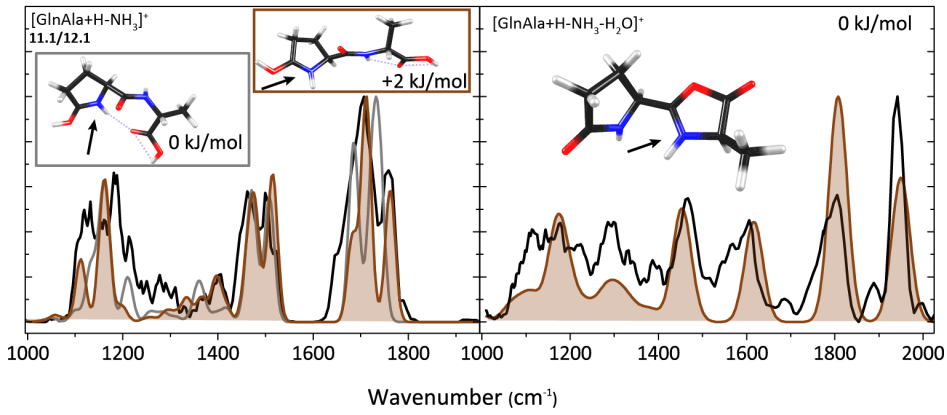
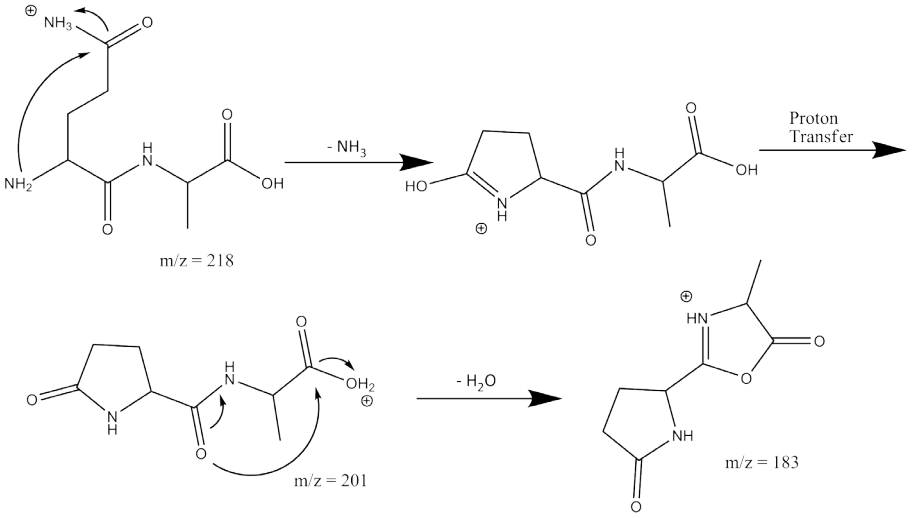


Figure 4-4. *(left) Measured IRMPD spectrum (black) of* $[GlnAla+H-NH_3]^+$ *compared with computed spectra (grey and shaded brown) for two nearly iso-energetic conformers of the lowest energy isomer* ***11.1*** *(see Table 4-4). (right) Measured spectrum for* $[GlnAla+H-NH_3-H_2O]^+$. *(black) along with the computed spectrum for the structure resulting from* H_2O *loss from the C-terminus. These spectra suggest that product ions are formed on the reaction path indicated in brown in Scheme 4-1 and more detailed in Scheme 4-5.*



Scheme 4-5. *Suggested reaction mechanism for the sequential loss of NH_3 and H_2O from protonated QA as derived from the product ion structures identified from their IRMPD spectra.*

4.4. Conclusion

The sensitivity and versatility of a commercial ion trap mass spectrometer combined with a powerful tunable IR laser and quantum-chemical calculations provides ample opportunities to apply ion spectroscopy for deep molecular structure probing in MS^n reaction networks, from which reaction mechanisms can be qualitatively inferred. In this study, we showcase these possibilities by constructing a reliable and comprehensive molecular structure map of the deamidation reaction mechanisms occurring in protonated peptides containing Asn and Gln residues (see Figure 4-5). These suggested mechanisms should form a useful starting point for computational modeling of the potential energy surfaces upon which the reactions proceed.[13,64-66]

Deamidation of $[AlaAsn+H]^+$ leads to a succinic anhydride structure, in contrast to what had been suggested for the deamidation of $[ValAsn+H]^+$ on the basis of computational investigation alone.[58] This reaction is also different from that occurring in solution, where succinimide-like structures are formed.[39-41] The mechanism established here for the peptide with Asn in the C-terminal position is analogous to that found for deamidation of the protonated Asn amino acid (chapter 3 and ref.[7]). Deamidation of $[AsnAla+H]^+$ proceeds along two bifurcating reaction pathways as evidenced by different spectral signatures observed in

different IR dissociation channels. Nucleophilic attack of the peptidyl amide oxygen on the side chain generates a structure featuring a 5-membered lactone ring. Alternatively, a succinimide product ion is formed upon nucleophilic attack by the peptidyl amide nitrogen onto the side chain carbonyl carbon. These mechanisms are in accord with recently established pathways for deamidation of protonated $[AsnGly+H]^+$.[59]

For $[AlaGln+H]^+$, the sequential loss of two NH_3 molecules is in agreement with earlier MS/MS studies[9], and here we add product ion structures and suggest reaction mechanisms. Elimination of NH_3 from the side-chain is induced by nucleophilic attack by the peptidyl amide nitrogen leading to a C-terminal pyroglutamic acid moiety. Subsequent elimination of NH_3 from the N-terminus leads to a structure with two fused 5-membered rings. For deamidation of $[GlnAla+H]^+$, formation of a pyroglutamyl residue at the N-terminus is established, in line with common knowledge for peptides with an N-terminal Gln residue.[2,3,39,42]

One thing that all reaction mechanisms have in common is the fact that deamidation occurs in all four dipeptides from the side chain amide group, which is also the case for the Gln and Asn amino acids in chapter 3. Moreover, all four dipeptides follow a reaction mechanism analogous to that in the respective amino acids, in the sense that the nucleophile driving the expulsion of an ammonia molecule is identical to the one in the amino acids. This is exemplified by deamidation of $[AlaAsn+H]^+$ and $[AsnAla+H]^+$, induced by nucleophilic attack of the backbone carbonyl oxygen of the Asn residue. Deamidation product ions from protonated $[AlaGln+H]^+$ and $[AsnGln+H]^+$ possess a pyro-glutamate-like structure induced by nucleophilic attack by the backbone nitrogen of the Gln residue, analogous to the mechanism in the glutamine amino acid. Only for $[AsnAla+H]^+$,, a parallel mechanism unlike that in the amino acid has been uncovered, involving the backbone nitrogen atom as the nucleophile.

The different side chain length of Asn and Gln implies that if $[AlaGln+H]^+$ followed the same mechanism as $[AlaAsn+H]^+$, and $[GlnAla+H]^+$ followed the same mechanism as $[AsnAla+H]^+$, the higher energy 6-membered ring structures **8** and **13** would have formed for $[AlaGln+H]^+$ and $[GlnAla+H]^+$, respectively (see Tables 4-3 and 4-4). Instead, the backbone amide nitrogen acts as nucleophile leading to lower energy 5-membered ring structures. Similarly, if this reaction would occur for $[AlaAsn+H]^+$ and $[AsnAla+H]^+$, it would lead to unfavorable 4-membered ring structures. The difference in side chain length thus has decisive influence on the deamidation reaction mechanisms: for Asn containing peptides, the nucleophile is

the carbonyl O-atom at the C-terminal side of the residue, while for Gln-containing peptides, the nucleophile is the backbone N-atom on the N-terminal side of the residue.

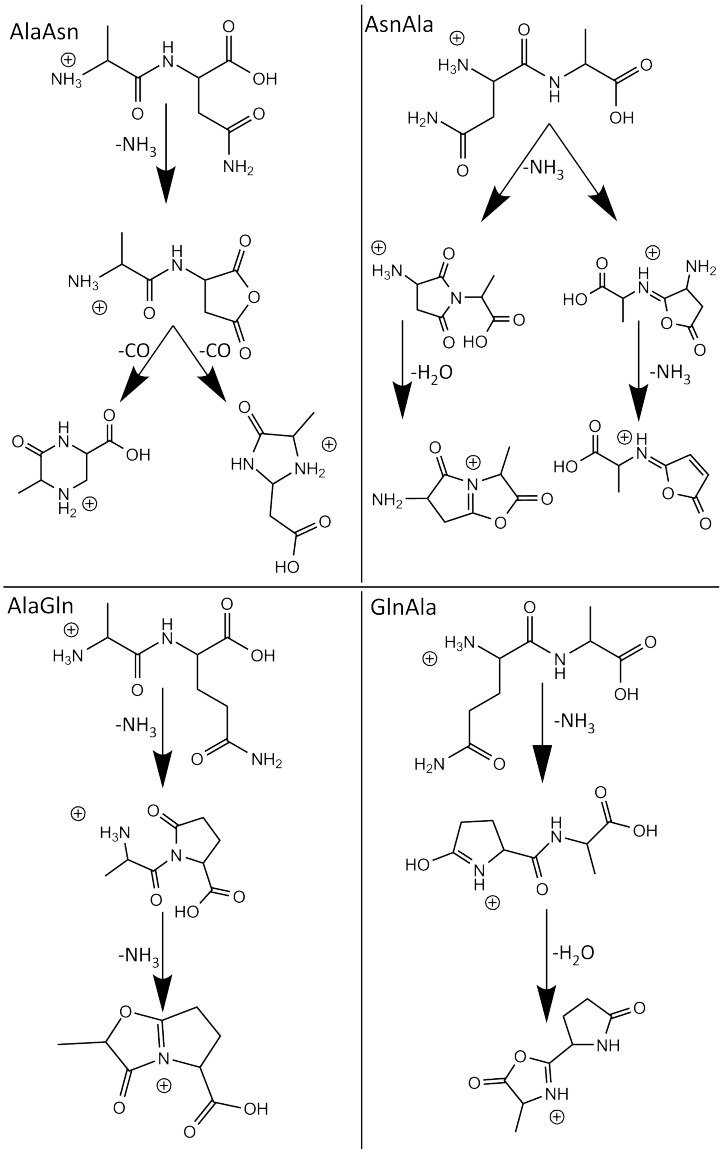


Figure 4-5. *Molecular structure map of deamidation reactions in the Asn and Gln containing dipeptides based on the fragment ion structures established on the basis of their IR spectra.*

Chapter 5

Deamidation of protonated asparagine-valine investigated by a combined spectroscopic, guided ion beam, and theoretical study

---

Peptide deamidation of asparaginyl residues is a spontaneous post-translational modification that is believed to play a role in aging and several diseases. It is also a well-known small-molecule loss channel in the MS/MS spectra of protonated peptides. Here we investigate the deamidation reaction, as well as other decomposition pathways, of the protonated dipeptide asparagine-valine ($[AsnVal+H]^+$) upon low-energy activation in a mass spectrometer. Using a combination of infrared ion spectroscopy, guided ion beam tandem mass spectrometry, and theoretical calculations, we have been able to identify product ion structures and determine the energetics and mechanisms for decomposition. Deamidation proceeds via ammonia loss from the asparagine side chain, initiated by a nucleophilic attack of the peptide bond oxygen on the γ-carbon of the Asn side chain. This leads to the formation of a furanone-ring containing product ion characterized by a threshold energy of 129 ± 5 kJ/mol (15 kJ/mol higher in energy than dehydration of $[AsnVal+H]^+$, the lowest energy dissociation channel available to the system). Competing formation of a succinimide-ring containing product, as has been observed for protonated asparagine-glycine ($[AsnGly+H]^+$) and asparagine-alanine ($[AsnAla+H]^+$), was not observed here. Quantum-chemical modelling of the reaction pathways confirms these subtle differences in dissociation behavior. Measured reaction thresholds are in agreement with predicted theoretical reaction energies computed at several levels of theory[*].

[*] Adapted from: Kempkes, L. J. M.; Boles, G. C.; Martens, J.; Berden, G.; Armentrout, P. B.; Oomens, J., Deamidation of Protonated Asparagine–Valine Investigated by a Combined Spectroscopic, Guided Ion Beam, and Theoretical Study. *The Journal of Physical Chemistry A* **2018,** *122* (9), 2424-2436.

5.1. Introduction

Various mass spectrometry (MS) studies have focused on the loss of small neutral molecules from protonated peptides containing asparagine (Asn) and glutamine (Gln).[1,3,7,9,38,58,59,63,67,68] Deamidation reactions of Asn and Gln have been the focus of significant attention, inspired by the spontaneous post-translational deamidation processes in proteins. These processes are believed to play a role in several diseases such as Alzheimer's and Parkinson's but also in aging effects for which it acts as a molecular clock.[4,5,39-41,69] Furthermore, the deamidation rates of several peptides containing Asn have been investigated and vary over three to four orders of magnitude.[70] The fastest deamidation rates were observed for peptides that contain smaller amino acid residues near Asn, presumably because they cause little to no steric hindrance to inhibit the deamidation process, although the detailed rationale for these variations has not yet been established.

The deamidation of Asn residues has been studied in condensed media and it was shown that the reaction proceeds via a succinimide structure [40,41]. Subsequent hydration leads to the formation of an aspartic acid (Asp) or iso-aspartic acid residue. More recently, the deamidation of $[Asn+H]^+$ was studied in the gas phase using a combination of threshold collision-induced dissociation (TCID) and high-level computations of the potential energy surface.[1] These results indicated that deamidation proceeds via a nucleophilic attack of the N-terminal nitrogen atom on the γ-carbon atom of Asn to form a succinic anhydride structure for $[Asn+H–NH_3]^+$. This identification was later confirmed by infrared action spectroscopy (see chapter 3).

Earlier gas-phase studies have also examined the deamidation reactions of protonated dipeptides that contain Asn, involving glycine as the second residue (AsnGly), valine as the first residue (ValAsn), and alanine as the first and second residue, (AlaAsn) and (AsnAla). The deamidation of $[AlaAsn+H]^+$ forms a succinic anhydride structure, initiated by a nucleophilic attack of the carbonyl oxygen on the Asn side chain (see chapter 4). $[AsnAla+H]^+$ follows a bifurcating mechanism leading to both a furanone (an imino-butyrolactone) and a succinimide structure. These reactions are initiated by the amide bond oxygen and the amide bond nitrogen, respectively. The deamidation of $[AsnGly+H]^+$ leads to a succinimide structure at threshold energies, with contribution from the furanone product at higher energies [59]. The deamidation reaction of $[ValAsn+H]^+$ has been investigated by theoretical modeling at the density functional theory (DFT) level [58]. According to this study, the nucleophilic attack occurs from the backbone amide oxygen onto

the side-chain carbonyl carbon, leading to a 6-membered lactone structure. Summarizing, ammonia loss takes place from the side chain in all cases considered thus far; however, differences in reaction rates and the mechanisms for deamidation as the adjacent residue varies remain incompletely understood.

In the present study, we investigate the deamidation of protonated asparagine-valine ($[AsnVal+H]^+$) and its MS^3 product ion. Valine contains an isopropyl side chain and is expected to have a larger steric hindrance in forming the deamidation product ion in comparison to alanine and glycine. The structures of the product ions were elucidated using a combination of infrared (IR) ion action spectroscopy, guided ion beam experiments, and quantum-chemical calculations.

Ion spectroscopy is a compelling technique for detailed structural characterization of small gas-phase ions in MS experiments. Using this technique, earlier studies showed it is possible to determine the structures of parent ions, as well as their reaction products after CID, on the basis of the ability to distinguish between isomers, protonation sites, and conformers.[25,30,44,46,71] For an accurate determination of thermochemistry involved during the fragmentation of peptides, TCID carried out in a guided ion beam tandem mass spectrometer has been applied.[59] Specifically, bond dissociation energies of metal-peptide interactions[72] as well as threshold energies for peptide fragmentation,[1,7,59,65,73] including deamidation reactions, have been established. Such information allows for the comparison of experimental and theoretical values such that likely decomposition mechanisms and product identities can be verified, a key aspect of TCID studies.

Separately, ion spectroscopy and guided ion beam experiments in combination with computational chemistry have already been proven as powerful techniques to establish the molecular structures of deamidated ions. The present study represents a novel combination of these techniques to assign the molecular structure of the product ions with high confidence. Additionally, complete quantum-chemical calculations were performed allowing for the identification of likely decomposition pathways. Such theoretical calculations were valuable in identifying key rate-limiting elementary steps of the possible reaction mechanisms for comparison with experimental TCID values. In all cases, the lowest energy pathways located for the observed fragmentation were in agreement with the experimental results.[1,7,59,65,73] Combining the structural and energetic information, the present study enables the identification of product ions and allows the differences in deamidation rates among different dipeptides to be explained.

5.2. Experimental and computational methods

5.2.1. Infrared ion spectroscopy

Infrared multiple photon dissociation (IRMPD) spectra of the protonated AsnVal precursor ion and the deamidated fragment ions are measured in a modified 3D quadrupole ion trap mass spectrometer (Bruker, AmaZon Speed ETD, Bremen, Germany) [60] coupled to the beamline of the infrared free electron laser FELIX [27]. The AsnVal sample was purchased from Biomatik (Canada) and used for all experiments without further purification. For electrospray ionization (ESI), peptides were dissolved in 50:50 acetontrile:water solution with addition of 0.1% formic acid. A concentration of 10^{-6} M was used to generate the protonated precursor peptide. MS^n fragments were produced by low-energy collisional activation with helium for 40 ms with an amplitude parameter of approximately 0.3 V. The isolated peptide fragment ions were irradiated with two 6 µs long macropulses from FELIX at a repetition rate of 10 Hz, each pulse having an energy of 5-40 mJ and a bandwidth of about 0.5% of the center frequency. IRMPD spectra were generated from the relative dissociation yield at each FELIX wavelength via $\sum I(fragment\ ions)/\sum I\ (fragment\ ions + precursor\ ion)$, where *I* is the ion intensity.[26] The dissociation yield at each laser frequency was determined from five averaged mass spectra and was linearly corrected for the frequency dependent variation the infrared laser pulse energy. All spectra were recorded over the 800 - 2000 cm^{-1} range.

5.2.2. Guided ion beam mass spectrometry

Kinetic energy dependent cross sections for the decomposition of $[AsnVal+H]^+$ with Xe were measured using a guided ion beam tandem mass spectrometer (GIBMS) that has been described in detail elsewhere.[74-76] Briefly, ions were generated using an ESI source[77] and 10^{-4} M AsnVal in 50:50 MeOH/H_2O solution, acidified with an appropriate amount of acetic acid, and syringe-pumped at a rate of 0.35 µL/hr into a 35 gauge stainless steel needle biased at 2200 – 2400 V relative to ground. Ions were directed through a heated capillary at 80 °C into a radio frequency (rf) ion funnel,[78] where they were focused into a tight beam. After exiting the ion funnel, the ions entered an rf trapping hexapole ion guide where the ions underwent on the order of 10^4 thermalizing collisions with ambient gas. As demonstrated in earlier studies, ions produced in the source region have a Maxwell-Boltzmann distribution of rovibrational states at 300 K.[72,77,79-82]

The precursor $[AsnVal+H]^+$ ions were extracted from the source and mass selected using a magnetic momentum analyzer, decelerated to a well-defined and variable kinetic energy, and focused into an rf octopole ion guide that traps the ions radially,[83,84] which minimizes losses of product and reactant ions. The ions passed through a collision cell containing xenon[85,86] at a sufficiently low pressure (≤ 0.4 mTorr) such that the opportunity for multiple collisions to occur was minimal, although the number of ions undergoing single collisions was sufficient to produce product cross sections characterized by a high intensity. (In this respect, the collisional activation of the two techniques is slightly different, where TCID uses single, higher energy collisions to induce dissociation whereas multiple, low energy collisions are employed in the ion spectroscopy studies.) The product and residual reactant ions drift to the end of the octopole guide, where they were extracted and focused into a quadrupole mass filter for mass analysis. High resolution conditions were used such that characterization of cross sections for channels separated by 1 amu was possible. As discussed previously,[59]high mass resolution conditions decrease the energy range suitable for TCID experiments, such that only high intensity channels were observed under these conditions.

Ions were detected with a high voltage dynode and scintillation detector,[87] and the signal was processed using standard pulse counting techniques. Ion intensities of reactants and products, measured as a function of collision energy, were converted to absolute cross sections as described previously.[74] Briefly, the calculation of the cross section from the ion intensities utilized a relationship that is directly analogous to the Beer-Lambert Law, specifically, $I = I_0 e^{-\rho\sigma l}$, where $\boldsymbol{I}$ is the reactant ion intensity after passing through the collision cell, $\boldsymbol{I}_0$ is the sum of the intensities of the reactant ion and all product ions, l is the length of the collision cell (8.3 cm), and ρ is the number density of the neutral reactant and equals $\boldsymbol{P/k_BT}$, where $\boldsymbol{P}$ and $\boldsymbol{T}$ are the pressure and temperature of the gas and $\boldsymbol{k_B}$ is Boltzmann's constant. The uncertainty in the relative cross sections is about ±5% and that for the absolute cross sections is about ±20%. The ion kinetic energy distribution was measured using a retarding potential analysis and found to be Gaussian with a typical full width at half maximum (FWHM) of 0.1 - 0.2 eV (lab). Uncertainties in the absolute energy scale are about ±0.05 eV (lab). Ion kinetic energies in the laboratory (lab) frame were converted to energies in the center-of-mass (CM) frame using $E_{CM} = E_{lab}\, m/(m + M)$, where M and m are the masses of the ionic and neutral reactants, respectively. All energies in this work are reported in the CM frame unless stated otherwise.

5.2.3. Cross section modeling

Thresholds of the TCID cross sections were modeled using equation [5-1],

$$\sigma_j(E) = \left(n\sigma_{0,j}/E\right)\sum_i g_i \int_{E_{0,j-E_i}}^{E}\left[k_j(E^*)/k_{tot}(E^*)\right]\left\{1 - e^{-k_{tot}(E^*)\tau}\right\}(E-\varepsilon)^{n-1}d\varepsilon \quad [5\text{-}1]$$

where $\sigma_{0,j}$ is an energy-independent scaling factor for channel *j*, *n* is an adjustable, empirical representation of factors that describe the efficiency of the energy transfer during collision and varies with the complexity of the system being studied,[75] *E* is the relative kinetic energy of the reactants, $E_{0,j}$ is the threshold for dissociation of the ground electronic and rovibrational state of the reactant ion at 0 K for channel *j*, τ is the experimental time for dissociation (~5 x 10^{-4} s, as measured by previous time-of-flight studies),[75] ε is the energy transferred from kinetic to internal modes during the collision, and *E** is the internal energy of the energized molecule (EM) after the collision, so that $E^* = \varepsilon + E_i$. The summation is over the rovibrational states of the reactant ions, *i*, where E_i is the excitation energy of each state and g_i is the fractional population of those states ($\Sigma g_i = 1$). The Beyer–Swinehart-Stein-Rabinovitch algorithm[88-90] was used to evaluate the number and density of the rovibrational states and the relative populations g_i were calculated for a Maxwell-Boltzmann distribution at 300 K. The term $k_j(E^*)$ is the unimolecular rate constant for dissociation of the EM to channel *j* via its rate-limiting transition state (TS). The rate coefficients $k_j(E^*)$ and $k_{tot}(E^*)$ are defined in Rice-Ramsperger Kassel-Marcus (RRKM) theory [91,92] as

$$k_{tot}(E^*) = \sum_j k_j(E^*) = \sum_j d_j N_j^{\dagger}(E^* - E_{0,j})/h\rho(E^*) \quad [5\text{-}2]$$

where d_j is the reaction degeneracy of channel *j*, $N_j^{\dagger}(E^* - E_{0,j})$ is the sum of rovibrational states for the TS of channel *j* at an energy $E^* - E_{0,j}$, and $\rho(E^*)$ is the density of states of the EM at the available energy *E**. These rate coefficients allow both kinetic shifts (where the probability of dissociation is given by the term $\left\{1 - e^{-k_{tot}(E^*)\tau}\right\}$ in eq (1) and competition between multiple channels (which is controlled by the ratio of rate coefficients in eq (1), $\left[k_j(E^*)/k_{tot}(E^*)\right]$) to be modeled accurately [93,94].

To evaluate the rate coefficients in equations [5-1] and [5-2], values for rovibrational energies for the EM and the rate-limiting TSs were determined from quantum-chemical calculations described below. Additionally, the entropy of activation at 1000 K for each dissociation channel was calculated as described in detail elsewhere.[93] The model cross sections of eq (1) were convoluted with the kinetic energy distributions of the reactants[74] and compared to the experimental

data. A nonlinear least-squares analysis was used to provide optimized values for $\sigma_{0,j}$, n, and $E_{0,j}$. The uncertainty in $E_{0,j}$ (reported as one standard deviation) was estimated from the range of threshold values determined from multiple sets of data, variations in the parameter n (±10% around the optimum value), variations in vibrational frequencies (±10%), changes in τ by factors of 2, and the uncertainty in the absolute energy scale (0.02 eV). Uncertainties associated with variations in the vibrational frequencies that control the internal energy and the kinetic shifts were assessed independently.

5.2.4. Computational chemistry

For all product ion structures, a first computational exploration was performed using DFT at the B3LYP/6-311+G(d,p) level of theory using Gaussian 09 revision D01.[31] Different protonation sites were considered for each structure. First guess structures were optimized and their computed infrared spectra were compared with the experimental one. For the lowest energy structure and the spectroscopically best matching structure, potential energy surfaces were further explored to identify the lowest energy conformers using a molecular mechanics/molecular dynamics (MM/MD) approach employing AMBER 12.[32] No MM/MD calculations were performed for the MS^3 fragments, as their conformational flexibility was small enough that chemical intuition was sufficient to generate input structures for DFT optimization. Within AMBER, an initial MM geometry optimization was performed, followed by a simulated annealing procedure up to 500 K resulting in 500 structures. These structures were grouped on the basis of their structural similarity using appropriate rms atom position criteria and yielded 20-30 candidate structures. These candidate structures were then optimized using DFT and their predicted IR spectra were compared with the experimental spectra. All computed harmonic vibrational frequencies were scaled by 0.975 and convoluted with a 25 cm^{-1} FWHM Gaussian line shape to facilitate comparison with experimental spectra. Finally, single-point 298 K Gibbs free energies were calculated from MP2(full)/6-311+G(2d,2p) electronic energies using the B3LYP/6-311+G(d,p) optimized structures and zero point vibrational energy (ZPE) corrections. The computational procedure is described in more detail elsewhere [15,38,60].

Independently, and to ensure that the ground structures (GS) of reactant and product species were correctly identified, structures analogous to those previously

reported for $[AsnGly+H]^+$ [59] were also used as starting structures. Here, optimizations of all unique low-energy structures were conducted at the B3LYP/6-311+G(d,p) level of theory. Starting structures for additional key reaction intermediates and all TSs were also initially taken from the resulting structures reported from the $[AsnGly+H]^+$ study [59]. Rotational constants and vibrational frequencies were calculated from optimized structures, and vibrational frequencies were scaled by a factor of 0.989 [95] when used for the determination of internal energy, RRKM calculations, and ZPE corrections. Single point energies of all reaction species were calculated using the 6-311+G(2d,2p) basis set at the B3LYP, B3P86, and MP2(full) levels using B3LYP/6-311+G(d,p) geometries and ZPE corrections.

5.3. Results and discussion

5.3.1. Precursor ion $[AsnVal+H]^+$

Figure 5-1 presents three low-energy conformers of the precursor ion $[AsnVal+H]^+$, with their relative energies calculated at the B3LYP, B3P86, and MP2(full)/6-311+G(2d,2p) levels of theory using B3LYP/6-311+G(d,p) geometries. Conformers are named according to their protonation sites including additional hydrogen bonds by using the designation [X,Y,Z] where X = protonated atom and Y/Z are sites that hydrogen bond to the protonated site, in order of increasing hydrogen bond length. The protonation site is followed by the series of dihedral angles starting from the N-terminal side-chain amide-group nitrogen to the C-terminal carboxylic acid. Backbone nitrogen and oxygen atoms are numbered by residue along the backbone chain starting from the N-terminus. Side-chain nitrogen and oxygen atoms are designated by a superscript s. Dihedral angles are distinguished as cis (c, for angles between 0°–45°), gauche (g, 45°–135°), or trans (t, 135°–180°).

The three lowest energy $[AsnVal+H]^+$ conformers are all located within 10 kJ/mol of each other at all levels of theory and protonate at the amine nitrogen (N^1) of the backbone. As shown in Figure 5-1, the N^1 protonation site interacts via hydrogen bonds with the carbonyl oxygen of both the side-chain amide (CO^s) and backbone amide (CO^1), [N^1,CO^s,CO^1], in all three conformations. B3LYP predicts the [N^1,CO^s,CO^1]-ttgtgtt conformer to be lowest in energy, whereas B3P86 and MP2 predict [N^1,CO^s,CO^1]-gggtgtt to be lowest in energy. These two structures differ in the hydrogen bonding interaction of the carbonyl group of the Val residue, CO^2: either with HN^2 in ttgtgtt or in a head-to-tail arrangement with HN^s in gggtgtt. The low-lying ttgtttt conformer lies only 2 – 3 kJ/mol above the similar ttgtgtt conformer, and they differ primarily in the orientation of the valine side chain.

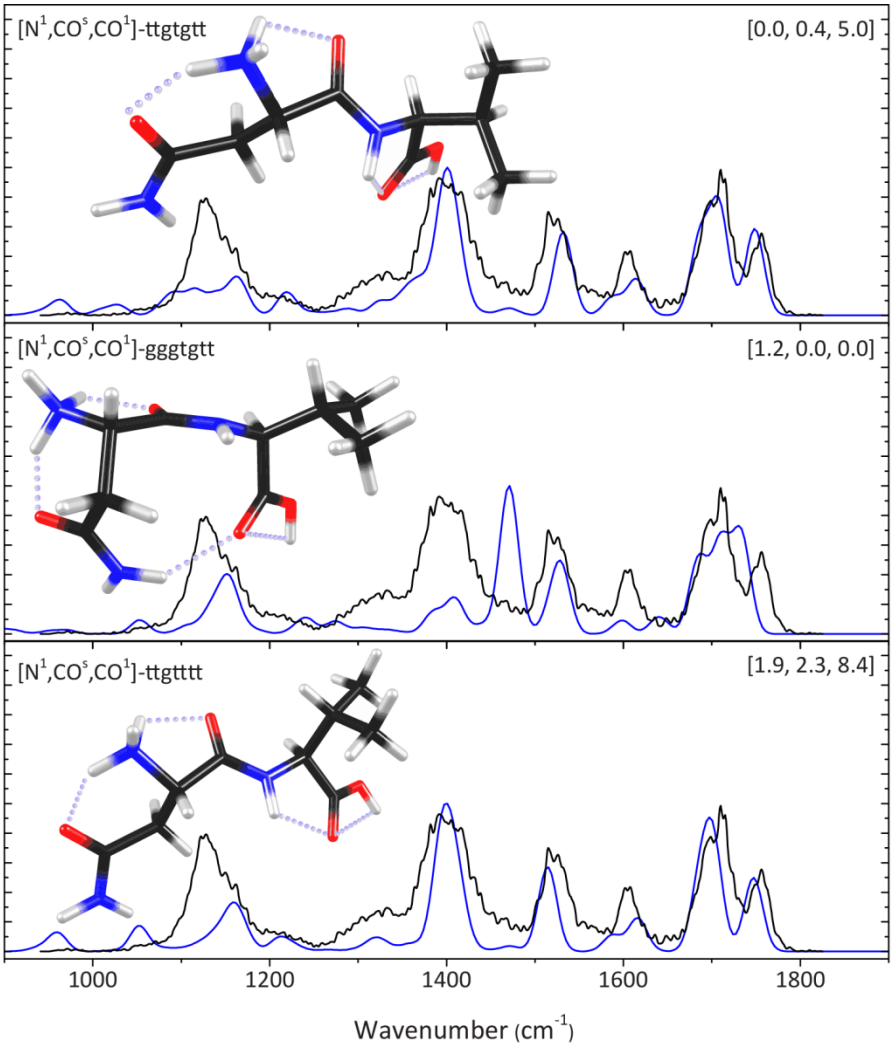


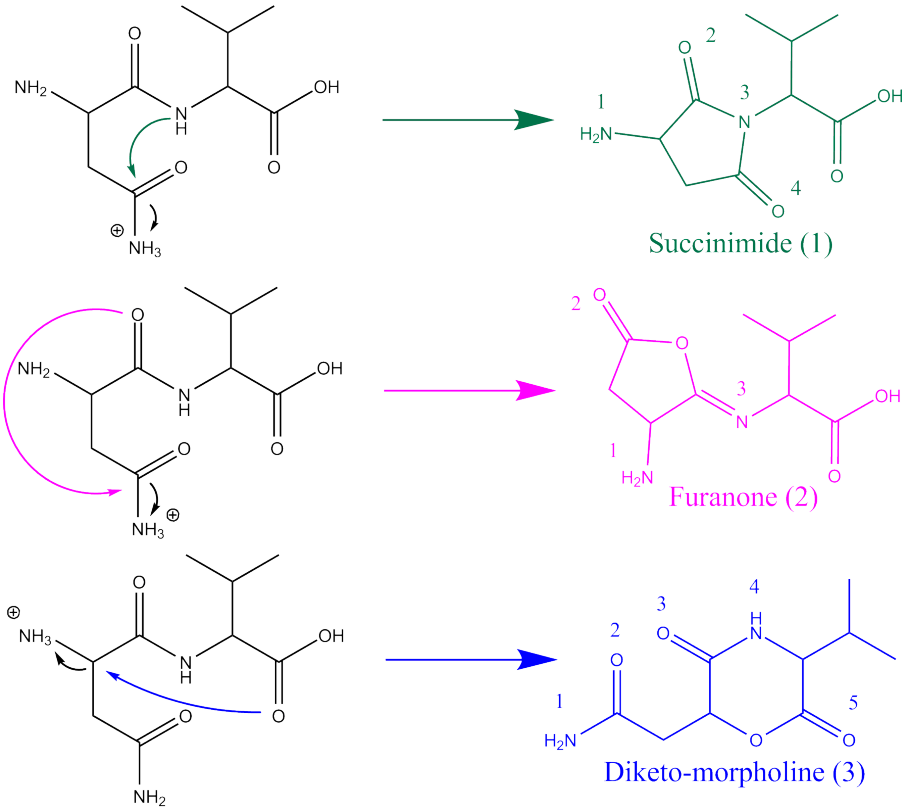
Figure 5-1. *Experimental IRMPD spectrum (black) of the $[AsnVal+H]^+$ precursor ion compared with the calculated spectra of the three lowest energy conformers (blue). All structures are protonated at the N-terminus. Relative 0 K single point energies (kJ/mol) at the B3LYP, B3P86, and MP2(full)/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) levels of theory are given in square brackets.*

To verify the conformation of the $[AsnVal+H]^+$ precursor ion, its IRMPD spectrum was recorded as shown in Figure 5-1 together with computed spectra for the three low-energy conformers of Figure 5-1. The top panel shows the calculated spectrum for the lowest energy conformer at the B3LYP level of theory (ttgtgtt), the middle panel that for the lowest energy conformer at the MP2 and B3P86 levels of theory (gggtgtt), and the bottom panel that for the conformer with an alternative Val side chain arrangement (ttgtttt). Clearly, the agreement in the 1300 - 1800 cm^{-1} range between experiment and the computation for the ttgtgtt conformer is convincing, although the computed spectrum for ttgtttt (bottom) is nearly as close. The peak at 1750 cm^{-1} is assigned to the C-terminal C=O stretch, and shifts to 1730 cm^{-1} for the gggtgtt conformer (middle). The band at 1710 cm^{-1} in all three computed spectra is assigned to the C=O stretch of the peptide bond. The shoulder at 1680 cm^{-1} in the middle spectrum results from the C=O stretch of the Asn side chain, which remains unresolved in the spectra of the other two conformers. The feature at 1610 cm^{-1} is assigned to NH bending of the NH_3 group in the top and the bottom panels, and shifts to 1640 cm^{-1} in the middle panel. The peak around 1530 (top and middle spectrum) and 1515 cm^{-1} (bottom spectrum) are assigned to NH bending of the amide bond.

The intense NH bending mode at 1470 cm^{-1} predicted for the gggtgtt conformer (middle) is not reproduced in the experimental spectrum, and clearly disqualifies this conformer. This mode is shifted to 1410 cm^{-1} in the other two conformers (top and bottom) in good agreement with observations. The main deviations between the computed spectra for ttgtgtt (top) and ttgtttt (bottom) is found between 1000 and 1200 cm^{-1}, where the sharper band in the bottom spectrum is assigned to OH bending at the C-terminus, while the weaker, broadened band in the top spectrum results from this OH bending mode in combination with several other non-localized bending modes. Clearly, none of the calculated spectra reproduce the intensity observed in the experimental spectrum at 1120 cm^{-1} well, although the ttgtgtt conformer does have a band at this position, whereas the ttgtttt conformer does not. Arguably perhaps, the top spectrum appears to be in better agreement with experiment and we conclude the protonated precursor ion to be predominantly in the $[N^1,CO^s,CO^1]$-ttgtgtt conformation although minor contributions from other conformers cannot be eliminated. Importantly, these results suggest that an open chained structure is preferred under our experimental conditions, whereas protein/peptide folding is more prominent in solution.

5.3.2. Deamidated Asparagine-Valine, $[AsnVal+H-NH_3]^+$

Loss of ammonia from $[AsnVal+H]^+$ can reasonably occur through either ammonia loss from the Asn carboxamide side chain or from the N-terminus (technically a deamination), especially considering the mobile proton model, which suggests that the proton can migrate from the most basic site to a less favored protonation site after collisional activation.[6,18-20,57] It is also not unreasonable to assume that ammonia loss is accompanied by a nucleophilic rearrangement involving the electrophilic carbon adjacent to the NH_3 leaving group.



Scheme 5-1. *Possible nucleophilic attack rearrangements leading to the deamidation of $[AsnVal+H]^+$. Loss of ammonia from the side chain (top and middle) and from the N-terminus (bottom) is considered. Numbered atoms are used to indicate different protonation sites in each of the product ion structures.*

Scheme 5-1 shows possible rearrangements leading to the expulsion of ammonia from $[AsnVal+H]^+$ with concomitant formation of a 5- or 6-membered ring structure. The green arrow indicates a nucleophilic attack of the amide bond

nitrogen onto the side chain, leading to a succinimide structure (**1**). The pink arrow indicates a nucleophilic attack of the amide bond oxygen onto the side chain leading to a furanone structure (**2**). The blue arrow indicates a nucleophilic attack of the C-terminal oxygen onto the N-terminus, leading to a diketo-morpholine structure (**3**). Possible protonation sites (1 – 5) for each of these three final structures are indicated in Scheme 5-1. The geometries of all these possible structures were optimized, their infrared spectra were calculated, and their relative energies are provided in Table 5-1.

Name →	**1 Succinimide**		**2 Furanone**		**3 Diketo-morpholine**	
$[H]^+$ site ↓	B3LYP	MP2	B3LYP	MP2	B3LYP	MP2
1	**0**	**0**	+ 91	+ 109	+ 57	+ 86
2	+ 46	+ 57	+ 82	+ 102	**+ 43**	**+ 76**
3	+ 117	+ 112	**+ 15**	**+ 40**	+ 54	+ 83
4	+ 37	+ 47			+ 118	+ 143
5					+ 87	+ 125

Table 5-1. *Relative 298 K Gibbs Free Energies (kJ/mol) for Possible Deamidation Product Ion Structures of $[AsnVal+H]^+$. The relative energies are calculated at the B3LYP/6-311+G(d,p) and MP2(full)/6-311+G(2d,2p) levels using B3LYP/6-311+G(d,p) geometries. The numbers in the first column refer to the different protonation sites indicated in each structure in Scheme 5-1. Entries in boldface indicate the lowest energy species for each type of product structure.*

The overall lowest energy structure is the succinimide structure (**1.1**), where the proton resides on the amino group that was the N-terminus. The lowest energy furanone structure (**2.3**) is protonated at the nitrogen of the peptide bond and lies 15 – 40 kJ/mol above **1.1**. The lowest energy diketo-morpholine lies 43 – 76 kJ/mol above **1.1** and is protonated on the side-chain carbonyl (**3.2**), partly because there is a strong hydrogen bond with a keto group of the ring structure (position 3).

In Figure 5-2, the experimental IRMPD spectrum (black) for the deamidation product ion $[AsnVal + H - NH_3]^+$ is displayed along with calculated spectra for the lowest-energy isomer **1.1** (bottom left panel) and for the furanone isomer **2.3** (top left panel). Figure 5-3 shows the spectral comparison with diketo-morpholine structure **3.2.** It is immediately clear that the furanone structure provides an excellent match to the experimental spectrum throughout the experimental region studied, even though its energy is substantially higher than that for the succinimide structure **1.1**. For the furanone structure, the diagnostic high-frequency band near 1950 cm^{-1} is assigned to the CO stretch of the carbonyl group.

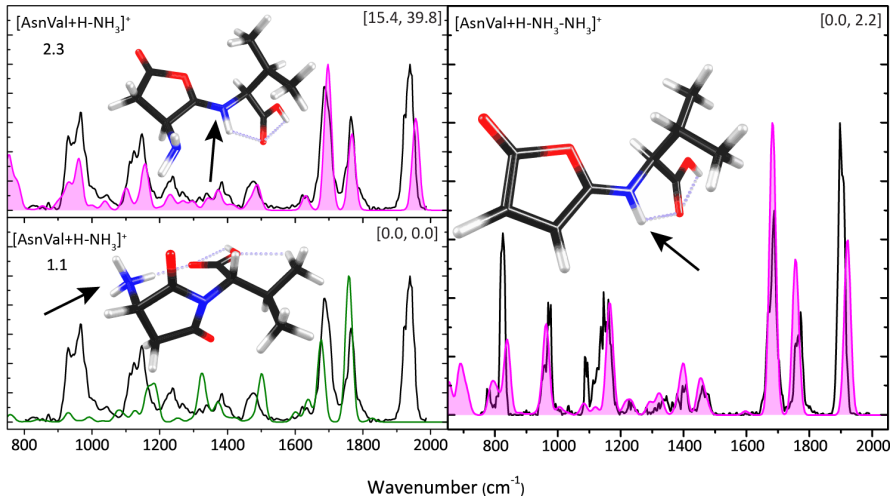


Figure 5-2. *IRMPD spectra (black) of $[AsnVal+H-NH_3]^+$ (left panels) and $[AsnVal+H-NH_3-NH_3]^+$ (right panel). The top left panel shows the computed spectrum for the furanone ring structure* ***2.3****, whereas the lower left panel shows that for the lowest-energy deamidated isomer, the succinimide structure* ***1.1****. In the right panel, the experimental spectrum is compared with the computed spectrum for the NH_3-loss fragment ion from the furanone ring structure. Protonation sites are indicated with arrows. Hydrogen bonds <2.5 Å are indicated with grey dashed lines. Energies (kJ/mol, in square brackets) are calculated at the B3LYP/6-311+G(d,p) (first value) and MP2(full)/6-311+G(2d,2p) levels using B3LYP/6-311+G(d,p) geometries (second value)*

The 1770 cm^{-1} band is the C-terminal C=O stretch (red-shifted because of the hydrogen bond to the protonated imine) and the band at 1710 cm^{-1} is the protonated imine CN stretch. The succinimide structure does not reproduce the high-frequency band nor bands below 1200 cm^{-1}. Here, the band predicted at 1760 cm^{-1} results from succinimide CO stretching, while the band at 1680 cm^{-1} is assigned to the C-terminal carbonyl stretch. Bands at 1640 and 1510 cm^{-1} are NH bending modes. Figure S3 in the S.I. of ref. [96] shows the calculated spectra of alternative conformers of structure **2.3** with higher energy. The main difference between the assigned structure and the higher energy conformers is the band around 900-1000 cm^{-1}, which is not reproduced well experimentally.

The right panel of Figure 5-2 shows the IRMPD spectrum of the MS^3 fragment ion (black) $[AsnVal+H-NH_3-NH_3]^+$ at *m/z* 198. Loss of an ammonia molecule from the furanone ring leads to the inlaid structure, for which the predicted spectrum is overlaid onto the experimental one. A reasonably good spectral match is found (especially above 1100 cm^{-1}), further supporting our assignment of the MS^2

fragment as the furanone structure. The spectrum computed for the furanone structure of $[AsnVal+H-NH_3-NH_3]^+$ exhibits main bands at 1910 cm^{-1} (C=O stretch of the furanone ring), 1770 cm^{-1} (C=O stretch of the C-terminal carboxyl group), and 1680 cm^{-1} (CN stretch of protonated imine). Most features at the low-frequency end of the spectrum, (750 – 1500 cm^{-1}) are reproduced reasonably well, namely the two sharp, intense bands at 825 and 960 cm^{-1} in the experimental spectrum, which correspond primarily to furanone CH out-of-plane bending and amide NH out-of-plane bending (calculated at 837 cm^{-1}) and furanone COC stretching, CNH bending, and furanone ring distortion (calculated at 961 cm^{-1}).

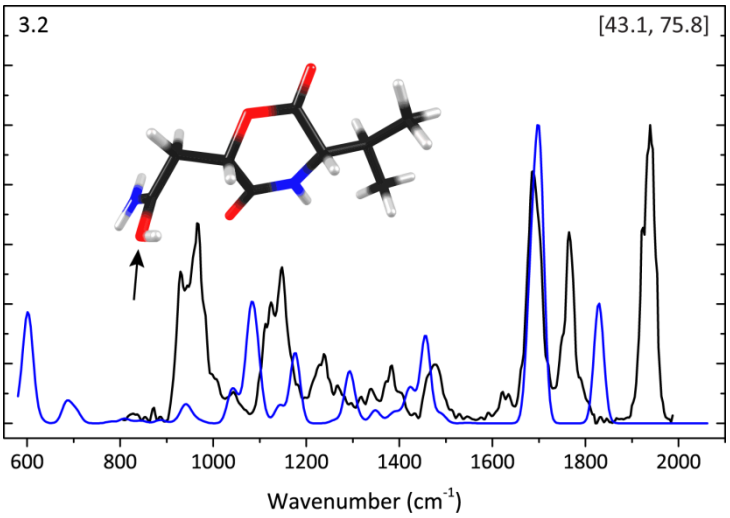


Figure 5-3. *IRMPD spectrum (black) of $[AsnVal+H-NH_3]^+$ compared to the computed spectrum of diketo-morpholine structure* ***3.2****. The protonation site is indicated with an arrow. Energies (kJ/mol, in square brackets) are calculated at the B3LYP/6-311+G(d,p) (first value) and MP2(full)/6-311+G(2d,2p) levels using B3LYP/6-311+G(d,p) geometries (second value).*

Figure 5-4 shows the comparison of the experimental spectrum of $[AsnVal+H-NH_3-NH_3]^+$ with other possible structures for additional loss of NH_3, none of which reproduce the experimental spectrum in its entirety (although the II-1 structures do match reasonably well in the high-frequency region, a result of similar CO and peptide bond CN stretching). Therefore, we assign the structure in the right panel of Fig. 5-2 on the basis of the good match across the entire spectral range.

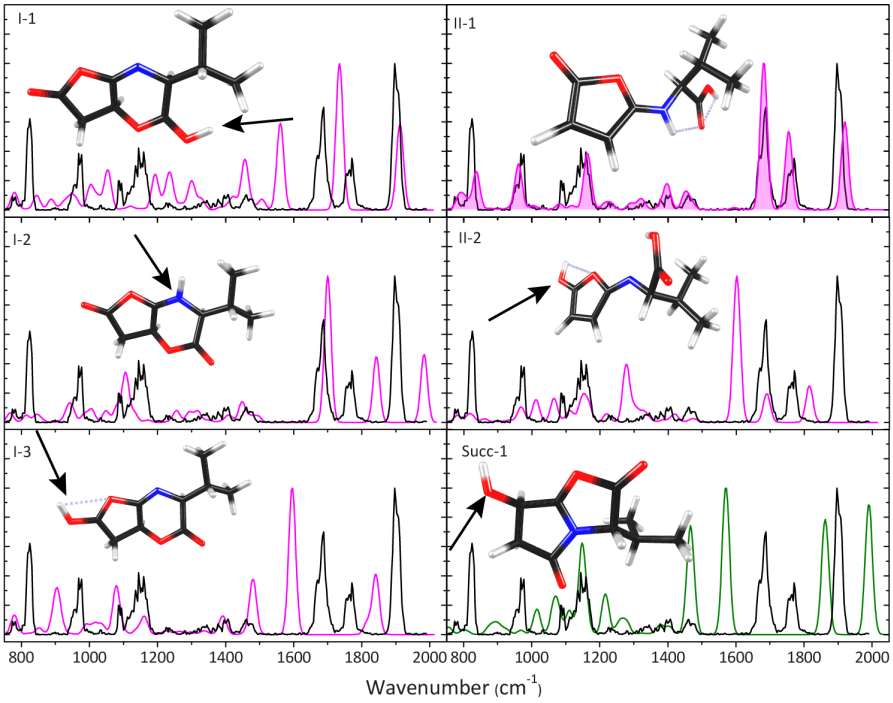


Figure 5-4. *IRMPD spectra (black) of [AsnVal+H-NH_3-NH_3]$^+$ compared to the computed spectra of NH_3 loss via a cyclization reaction (pathway I) or via NH_3 cleavage pathway (II). The protonation sites are indicated with an arrow. Energies (kJ/mol, in square brackets) are calculated at the B3LYP/6-311+G(d,p) (first value) and MP2(full)/6-311+G(2d,2p) levels using B3LYP/6-311+G(d,p) geometries (second value). To double check, the spectrum of the product from NH_3 loss from the succinimide was also calculated (see bottom right panel); the obvious mismatch reinforces the our structural assignment of [AsnVal+H-NH_3]$^+$ as a furanone ring containing ion.*

5.3.3. Cross sections for collision induced dissociation of [AsnVal+H]$^+$

Figure 5-5 shows the cross sections for the interaction of [AsnVal+H]$^+$ with Xe under high mass resolution conditions. Deamidation (m/z 215) and dehydration (m/z 214) are the dominant reaction channels, which is consistent with the dissociations observed previously for [AsnGly+H]$^+$ and [AsnAla+H]$^+$ [59,63]. At thresholds of about 0.8 eV higher, the formation of m/z 118 and 187 occur competitively, with m/z 87 and 198 arising still 0.3 eV higher. At even higher energies, the formation of m/z 115 and 173 is observed.

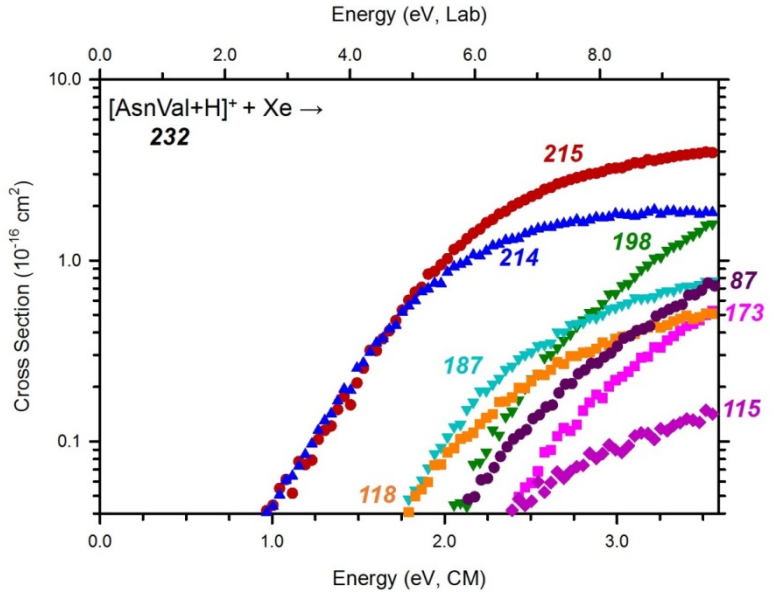
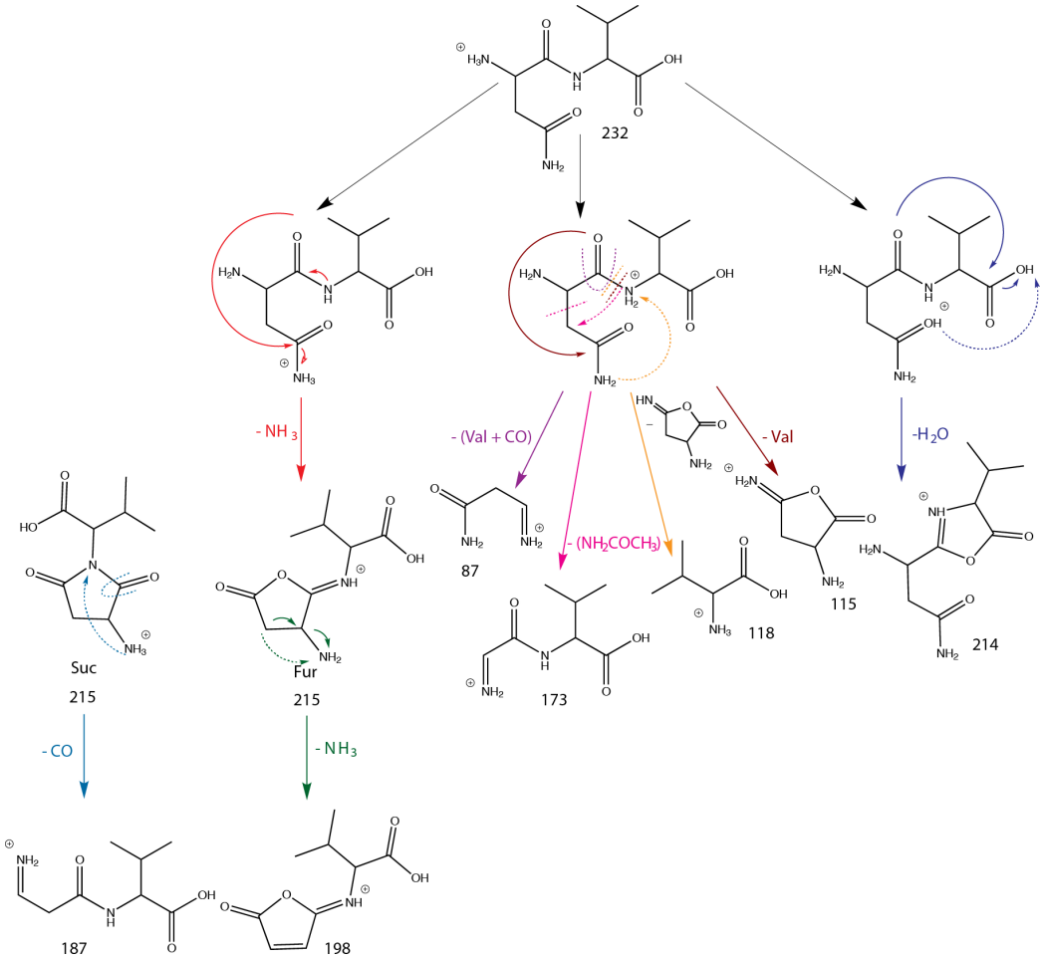


Figure 5-5. *Cross sections for the collision-induced dissociation of [AsnVal+H]$^+$ with Xe as a function of kinetic energy in the center-of-mass frame (lower x-axis) and laboratory frame (upper x-axis). Numbers indicate the mass to charge ratio of the ionic reaction products. Cross sections were measured at a pressure of 0.2 mTorr of Xe.*

As shown in Scheme 5-2, six primary channels are observed in the decomposition of [AsnVal+H]$^+$. Here, deamidation forms m/z 215, which is a protonated furanone species as identified in the ion spectroscopy experiment described above. Dehydration of [AsnVal+H]$^+$ leads to the b2-sequence ion, which is proposed to possess an oxazolone structure, similar to that observed in [AsnGly+H]$^+$ and [AsnAla+H]$^+$ studies.[59,63] Concomitant loss of (Val + CO) leads to the formation of m/z = 87, with analogous channels observed in the previous studies. Interestingly, parallel pathways are observed for the primary loss of Val (leading to m/z = 115) as well as the primary formation of [Val+H]$^+$ (m/z = 118), which deviates from the behavior exhibited by [AsnGly+H]$^+$ and [AsnAla+H]$^+$. Specifically, no analogous parallel dissociation was observed in the [AsnGly+H]$^+$ study[59], and in the analysis of [AsnAla+H]$^+$, only the loss of Ala was observed. The difference in behavior is presumably a reflection of the increasing complexity of the second residue, which influences the resulting relative proton affinities (PA) (where Gly < Ala < Val). Here, [Val+H]$^+$ formation is characterized by a lower threshold and greater intensity than the parallel loss of neutral Val across the entire energy range. Two secondary losses, m/z 198 and m/z 187, are assigned as sequential NH3 and CO losses from

the deamidation product, respectively. The structure of the m/z 198 product is as assigned from the ion spectroscopy experiment above. The m/z 173 product corresponds to the loss of the Asn side chain, leading to the product given in Scheme 5-1. An alternative cyclic structure (as presented in the $[AsnGly+H]^+$ study[59]) was also explored there, but found to be ~10 kJ/mol higher in energy.



Scheme 5-2. *Major decomposition reaction pathways for $[AsnVal+H]^+$ observed in the TCID analysis under high mass resolution conditions. Numbers indicate mass to charge ratios. Dashed lines indicate cleavages and dashed arrows indicate proton movement. Longer dashed lines indicate bond cleavages. Solid lines indicate bond formation. An alternative succinimide deamidation product is shown on the far left.*

5.3.4. Theoretical results for the mechanism of $[AsnVal+H]^+$ deamidation

The pathway yielding the furanone product ion parallels those located in the $[AsnGly+H]^+$ study.[59] Specifically, proton transfer from the protonated N-terminus to the side-chain amide nitrogen allows for the formation of the NH_3 leaving group. Subsequent C-NH_3 bond rupture is facilitated by nucleophilic attack of the backbone carbonyl of the first residue and occurs via $TS_{N\text{-}FUR}$ (132, 139, and 129 kJ/mol above the $[AsnVal+H]^+$ ground structure at the B3LYP, B3P86, and MP2 levels of theory, respectively). This leads to formation of furanone **2.3** (Table 5-1) as shown in Figure 5-6. Here, $TS_{N\text{-}FUR}$ is rate-limiting at the B3LYP and B3P86 levels of theory by 3 – 8 kJ/mol, whereas the MP2 level of theory instead predicts a product limited pathway where the furanone **2.3** + NH_3 product asymptote is 8.2 kJ/mol higher in energy than $TS_{N\text{-}FUR}$. Given this small energetic difference, the tight $TS_{N\text{-}FUR}$ is still likely to control the rate of reaction at the MP2 level of theory.

An alternative pathway for deamidation of the complex was also explored. Here, the rate-limiting step for succinimide formation, $TS_{N\text{-}Suc}$ (shown Figure S1 of ref.[96] and formed via a pathway parallel to that reported in a previous dipeptide analysis[59]), was found to be 40, 25, and 7 kJ/mol higher in energy than $TS_{N\text{-}FUR}$ at the B3LYP, B3P86, and MP2 levels of theory, respectively (see Table S1 in ref.[96]). This explains why the furanone **2.3** structure is formed even though the final succinimide **1.1** product is much lower in energy, Table 5-1. In the GIBMS studies, the deamidation channel is likely dominated by furanone formation at threshold energies, although contribution of both dissociation pathways is possible at higher energies. Contributions of both furanone and succinimide formation at higher energies have been reported previously in the analysis of $[AsnGly+H]^+$.[59]

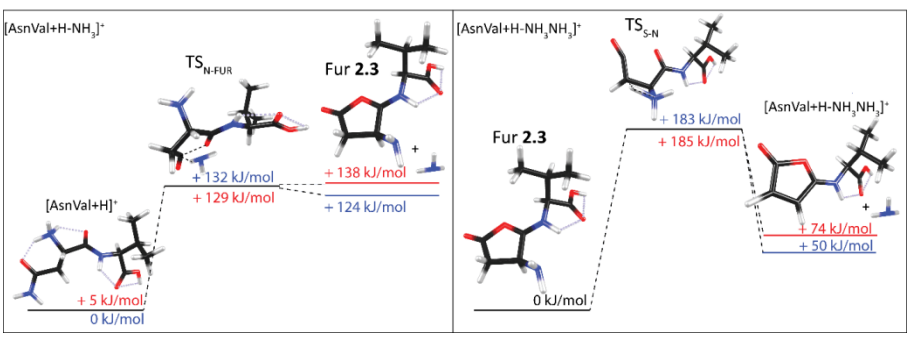


Figure 5-6. *Overview of the transition states for the primary and secondary deamidation reactions of $[AsnVal+H]^+$. Blue numbers indicate the 0 K B3LYP energies (kJ/mol), red numbers indicate 0 K MP2 energies (kJ/mol). Dotted lines indicate hydrogen bonds <2.5 Å whereas dashed lines indicate bonds being formed or broken in the transition states.*

5.3.5. Theoretical results for the mechanism of [AsnVal+H]+ dehydration

Similar to deamidation, dehydration of $[AsnVal+H]^+$ follows the mechanisms reported previously for $[AsnGly+H]^+$.[59] Namely, a proton transfer from the N-terminus to the carbonyl oxygen of the side-chain carboxamide (CO^5) is followed by dihedral angle rotations such that the complex arranges itself in a head-to-tail conformation. This orientation forms a hydrogen bond between the protonation site and carboxylic acid hydroxyl group (O^3H), Scheme 5-2. From here, a proton transfer from CO^5 to O^3H results in the formation of the H_2O leaving group, and nucleophilic attack of the CO^1 carbonyl initiates bond rupture of $C\text{-}O^3$ and oxazolone ring formation. This process (via TS_O, 103 – 118 kJ/mol above the GS) is rate-limiting at all levels of theory, where the oxazolone product lies 18 – 30 kJ/mol lower in energy than TS_O. Dehydration of $[AsnVal+H]^+$ is the lowest energy dissociation process observed, consistent with predicted reaction energies that are 16 – 35 kJ/mol lower in energy than the rate-limiting steps for the above-mentioned deamidation process.

5.3.6. Theoretical results for the formation of *m/z* 198 (-NH_3), *m/z* 197 (-H_2O), and *m/z* 187 (-CO) from the deamidation product

The deamidation product is observed to undergo sequential loss of another ammonia molecule (see Figures 5-2 - 5-5 and Scheme 5-2) and, although not observed experimentally, sequential loss of H_2O is also conceivable. Both reaction pathways were explored theoretically. Secondary ammonia loss from the Fur **2.3** product follows a straightforward mechanism that parallels that detailed in a forthcoming study on $[AsnAla+H]^+$. Here, proton transfer from the furanone CH_2 group to N^1 via $TS_{S\text{-}N}$ (where *S-N* stands for the *s*equential loss of an NH_3 group) results in the formation of the NH_3 leaving group. This process partially opens the furanone ring, where $C\text{-}NH_3$ bond rupture in the following step allows for ring closure. Notably, these motions are not concerted, such that $C\text{-}NH_3$ bond rupture via rate-limiting $TS_{S\text{-}N}$ (307 – 323 kJ/mol above the reactant ground structure) is ~50 kJ/mol higher in energy than ring closure, which occurs in the following step. These motions lead to the formation of $[AsnVal+H\text{-}NH_3\text{-}NH_3]^+$ as shown in Figure 5-2, characterized by a product asymptote of 174 – 212 kJ/mol.

At higher energies, a pathway for the sequential loss of water from Fur **2.3** was found in the computations. Here, proton transfer from N^1 to O^3H occurs resulting in dehydration and $N^1\text{-}CO^2$ bond formation, stabilized by the formation of bicyclic *m/z* 197 as shown in Figure S1 in the supplementary material of ref.[96]. Similar to secondary ammonia loss, $TS_{S\text{-}O}$ (360 – 367 kJ/mol above the ground state) is

predicted to be rate-limiting, with the *m/z* 197 + H_2O product asymptote located 170 – 188 kJ/mol lower in energy than $TS_{S\text{-}O}$. Note that $TS_{S\text{-}O}$ is >40 kJ/mol higher in energy than $TS_{S\text{-}N}$, plausibly explaining why subsequent loss of H_2O from Fur **2.3** is not observed in the IRMPD or TCID experiments.

The loss of neutral CO forming *m/z* 187 was also observed in the TCID experiments as a decomposition channel from the *m/z* 215 deamidation product with a threshold lower than loss of a second NH_3 group. Initially, a pathway exploring this decarbonylation of the furanone product was explored. Here, proton transfer to the NH_2 amino group (from the adjacent carbon) weakens the furanone ring, which can undergo ring opening in the following rate-limiting step, $TS_{S(FUR)\text{-}C}$ (362 – 389 kJ/mol above the GS). Once past this TS, the system eliminates CO. Notably, this process is considerably higher in energy than predicted reaction energetics for secondary ammonia loss, in contrast to the relative thresholds observed in the TCID cross sections.

Thus, decarbonylation of the alternative succinimide **1.1** product (which may be present at higher energies as discussed above) was also explored. Here, proton transfer from the NH_3^+ protonation site to the cyclic nitrogen weakens the succinimide ring. In the following step elongation of the N-C bond breaks the ring via $TS_{S(SUC)\text{-}C}$ (266 – 293 kJ/mol above the GS). Importantly, this process is 92 – 119 kJ/mol lower in energy than decarbonylation of the furanone **2.3** product, and 33 – 57 kJ/mol below NH_3 loss from furanone **2.3**. The latter is consistent with the cross sectional behavior observed in the TCID studies at threshold energies.

5.3.7. TCID data analysis and modeling

Equation 1 was used to analyze the thresholds for the primary competitive deamidation and dehydration channels in the decomposition of $[AsnVal+H]^+$, where the simultaneous modeling of these primary channels accounts directly for competition. Using parameters given in Table 5-2, the data were reproduced over the full range of the energy and cross section magnitude, as shown in Figure 5-7. The experimental deamidation channel shown includes the summation of the sequential dissociations forming *m/z* 187 and 197 (thus incorporating all relevant cross sections). For the reactions limited by a tight TS (deamidation via $TS_{N\text{-}FUR}$ and dehydration via TS_O), the TS frequencies used for the cross section modeling were taken from the theoretical results discussed above. Higher energy channels observed in Figure 5-5, including the sequential dissociations into channels *m/z* 187 and 197, could not be modeled accurately because their high energy thresholds lead to relatively low intensities.

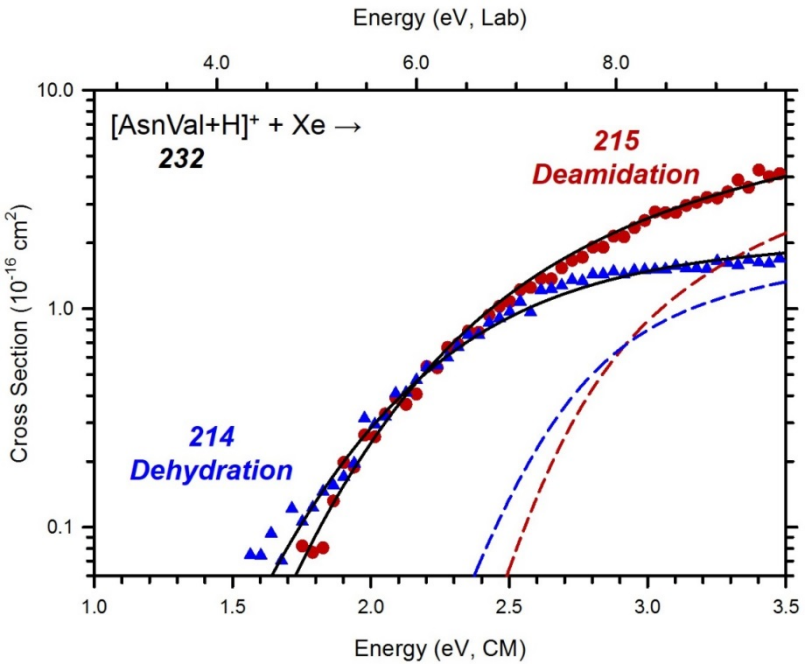


Figure 5-7. *Cross section models of the main decomposition products of [AsnVal+H]$^+$ as a function of collision energy with Xe in the center-of-mass frame (lower x-axis) and the laboratory frame (upper x-axis). Solid lines show the best fit to the data extrapolated to zero pressure (solid symbols), using the model of eq 1 convoluted over the neutral and ion kinetic and internal energy distributions. Dashed lines show the model cross sections in the absence of experimental kinetic energy broadening for reactants with an internal energy of 0 K.*

As shown in Figure S5 of the S.I. in ref. [96], zero pressure extrapolated deamidation and dehydration cross sections could be modeled over the full energy range fairly well using $TS_{N\text{-}FUR}$ and TS_O molecular parameters, respectively. However, improved fits in the threshold regions of these respective cross sections could be obtained by scaling the low-frequency modes (<900 cm^{-1}) of either $TS_{N\text{-}FUR}$ or TS_O (as shown in Figure 5-7). Here, low-frequency modes for TS_O were tightened by 2-5% while holding those for the tight $TS_{N\text{-}FUR}$ constant. (Notably, a comparable fit can be achieved by loosening the low-frequency modes of $TS_{N\text{-}FUR}$ by the same magnitude. In the current analysis, however, tightening of the TS_O modes is more appropriate because this approach does not alter the dominant dissociation pathway.) This analysis suggests that the calculated harmonic frequencies for TS_O provide a characterization of this channel that is slightly too loose relative to those for $TS_{N\text{-}FUR}$. Similar (and even larger) frequency scaling factors have been used in previous

studies on comparable systems.[7,59,71] Notably, although the analysis with no scaling (as given in Figure S5 in ref. [96]) provides a similar global representation of the data that is not quite as good in the critical threshold region, it results in threshold energies within the stated uncertainties of the values provided in Table 5-2. Additionally, the deamidation channel was only modeled using tight TS parameters, even though the MP2 level predicts a product limited pathway by 8 kJ/mol. However, as noted above, this small energy difference means that the tight TS barrier is still likely to control the reaction. Previous modeling analyses of similarly characterized pathways confirm this behavior [59]. It can be noted that the entropies of activation, $\Delta S^{\ddagger}_{1000}$, for these two channels indicate that dehydration is entropically disfavored compared with deamidation, which explains why the former cross section is smaller at higher energies even though it is less endothermic.

Reaction	TS	Frequency Scaling[a]	σ_0	n	E_0 (eV)	$\Delta S^{\ddagger}_{1000}$ (J/K mol)
Dehydration m/z 215	TS_O	1.02	15.1 (1.3)	1.2 (0.2)	1.18 (0.05)	-39 (1)
Deamidation[b] m/z 214	$TS_{N\text{-}FUR}$	1.00	15.1 (1.3)	1.2 (0.2)	1.18 (0.05)	-2 (1)

Table 5-2. *Fitting Parameters of Equation 1, Threshold Energies at 0 K, and Entropies of Activation at 1000 K for Cross Sections of the Reactions Indicated. Uncertainties are indicated in the parentheses.*

[a] *Frequency scaling factor applied to frequencies < 900 cm^{-1}.*

[b] *Values obtained using tight TS parameters, see text.*

For comparison of the thermodynamic information obtained experimentally to room temperature conditions, Table 5-3 provides the conversion from 0 K thresholds to 298 K enthalpies and Gibbs free energies. This conversion was accomplished using the rigid rotor/harmonic oscillator approximation with rotational constants and vibrational frequencies calculated at the B3LYP/6-311+G(d,p) level. Uncertainties listed were determined by scaling the vibrational frequencies by ± 10%.

Reaction	ΔH_0	$\Delta H_{298}-\Delta H_0$ [a]	ΔH_{298}	$T\Delta S_{298}$ [a]	ΔG_{298}
Dehydration, TS_O	114 (5)	0.4 (0.2)	114 (4)	-11.0 (0.1)	125 (4)
Deamidation, $TS_{N\text{-}FUR}$	129 (5)	3.7 (0.1)	133 (4)	-1.1 (0.2)	134 (5)

Table 5-3. *Enthalpies and Gibbs Free Energies of Reaction at 0 and 298 K for Deamidation and Dehydration Reactions. The parentheses indicate the uncertainties.*

[a] *Calculated using standard formulae and molecular constants at the B3LYP/6-311+G(d,p) level.*

5.3.8. TCID experimental versus theoretical results

Experimental threshold energies for the primary deamidation and dehydration reactions are compared with calculated theoretical reaction energies in Table 5-4 at several levels of theory. Such comparisons are important in ensuring that thermal ions were produced experimentally and can confirm the validity of the reaction mechanism. Here, "hot" ions would systematically result in threshold energies that are uncharacteristically low. For the deamidation pathway leading to furanone **2.3** formation, the experimentally determined threshold was found to be 129 $\pm$ 5 kJ/mol, in excellent agreement with B3LYP and MP2 theoretical values. B3P86 suggests a reaction threshold higher by about 10 kJ/mol. The competitive dehydration channel was found to have an experimental threshold of 114 $\pm$ 5 kJ/mol, which is within the uncertainty of the experimental value for the B3LYP and B3P86, whereas MP2 suggests a threshold lower by 11 kJ/mol (just outside two standard deviations).

For both decomposition processes, there is good agreement between experiment and theory when modeled competitively using parameters for the tight TS limited pathways. Such agreement is quantified via mean absolute deviations (MADs) from the experimental results as shown in Table 5-4, where all MADs are lower than 10 kJ/mol. B3LYP exhibits the lowest MAD at 3 kJ/mol, whereas B3P86 overestimates the reaction thresholds such that this level resulted in the highest MAD, 7 kJ/mol. The consistent agreement between theory and experiment for both modeled channels verifies that the necessary experimental conditions to produce thermal ions were used and confirms that the mechanisms located for these reactions are reasonable pathways. Notably, the experimental threshold for deamidation is not

consistent with predicted reaction energies found for succinimide formation, 145 – 172 kJ/mol, Table 5-5, consistent with the failure to observe this species in the ion spectroscopy experiments.

Reaction	TS	Experiment	B3LYP	B3P86	MP2(full)
Dehydration m/z 172,	TS_O	114 (5)	116	118	103
Deamidation m/z 173,	$TS_{N\text{-}FUR}$	129 (5)	132	139	129
MADs[a]			3 (1)	7 (4)	6 (8)

Table 5-4. *Experimental and Theoretical Reaction Energies (kJ/mol) for Decomposition Reactions of [AsnVal+H]*$^+$***. Uncertainties are indicated in the parentheses. The Calculations are performed at the stated level of theory using a 6-311+G(2d,2p) basis set with geometries calculated at B3LYP/6-311+G(d,p) level, ZPE corrections included.***

[c] *Calculated mean absolute deviations (MADs) from experiment for the deamidation and dehydration channels*.

5.3.9. Comparison to solution phase reactivity

Notably, the behavior observed in the current analysis (specifically formation of furanone, rather than succinimide) differs from solution phase studies.[40,41] On this basis, we explored solvated reaction pathways, both implicitly by optimizing reaction TSs and intermediates in a polarizable continuum model (PCM) cavity, as well as explicitly by the addition of one water molecule. Here, succinimide formation via a tetrahedral intermediate is favored over furanone formation in solution by 19 and 37 kJ/mol at the B3P86 and MP2 levels of theory, respectively, although B3LYP predicts succinimide is higher in energy than furanone formation by 18 kJ/mol. However, the better correlation between experiment and theory at the MP2 level of theory for the deamidation processes (as given below in Table 5-5) suggests that this level of theory likely handles the reaction dynamics more accurately. Notably, the MP2 level predicts the TS is ~60 kJ/mol lower in energy than the B3LYP level, an energy difference not likely to be water mediated to the degree that succinimide would be favored over furanone formation at the B3LYP level of theory. Thus, the differences between gas-phase and condensed-phase dynamics appear to be controlling the observed deamidation product in these respective studies.

Analysis of the TS motions involved in the solvated reactions for furanone and succinimide formation provides insight as to how the reaction dynamics are affected in gas-phase and condensed-phase media. Specifically, succinimide formation via tetrahedral intermediates proceeds through a rate-limiting TS characterized by several concerted proton tranfers.[9] These motions can be facilitated through by water mediation effects that lower the reaction barriers. However, furanone formation in solution proceeds through an analogous pathway as depicted in Figure 5-4, where concerted motions of bond rupture and bond formation are observed. Here, no water mediation effects are observed because the available water molecule does not facilitate the reaction, but rather stabilizes the NH_3 group throughout the reaction, leading to an increase of the reaction barrier. Thus, succinimide formation is most commonly reported, where subsequent hydrolysis yields a combination of iso-aspartic acid and aspartic acid. Although furanone formation is not likely energetically competitive with succinimide formation in condensed media, hydrolysis of such furanone structures could similarly lead to ring opening in solution, forming an Asp residue, but not the alternative iso-Asp.

5.3.10. Comparison to $[AsnGly+H]^+$ and $[AsnAla+H]^+$

In previous studies, our groups have investigated the deamidation reactions of protonated AsnGly and AsnAla. (ref. [59] and chapter 4) Interestingly, increasing the size of the C-terminal residue side-chain influences the reaction energetics for succinimide and furanone formation upon deamidation, see Table 5-5. For succinimide formation, a larger C-terminal residue side-chain correlates with an increase in the energy of the rate-limiting TS. Relative to $[AsnGly+H]^+$, rate-limiting TSs of the Ala and Val-containing species exhibit increases in energy of 12 and 21 kJ/mol at the B3LYP and B3P86 levels (4 and 12 kJ/mol at the MP2 level), respectively. Notably, succinimide formation occurs via parallel pathways in each of the complexes such that the difference in energy among these three systems appears to be a direct effect of increased steric hindrance. Conversely, for furanone formation, no clear correlation between C-terminal residue side chain size and TS energy is observed. At the B3LYP and B3P86 levels, specifically, reaction energies are within 1 kJ/mol of each other for the three dipeptide systems.

Experimentally, deamidation of $[AsnGly+H]^+$ and $[AsnAla+H]^+$ were observed to exhibit bifurcating reaction pathways, forming the succinimide and furanone product ions in parallel. For the $[AsnVal+H]^+$ dipeptide studied here, this is not the

case: only the furanone product is observed, in agreement with the trends in the computed reaction barriers.

Dehydration of all three systems was observed to compete with deamidation, in all cases being lower in energy but entropically disfavored. It can be seen that there is no systematic trend in the experimental dehydration energies as a function of side-chain complexity, but the experimental values match the trends predicted at the MP2 level nicely. Because the oxazolone formation occurs within the second residue containing the aliphatic side chains, one can imagine that a combination of inductive and steric effects could lead to the changes observed.

Reaction	Species	Experiment	B3LYP	B3P86	MP2(full)
Succinimide Formation	$[AsnGly+H]^+$	129 (6)	151	143	135
	$[AsnAla+H]^+$		163	155	137
	$[AsnVal+H]^+$		172	164	145
Furanone Formation	$[AsnGly+H]^+$		132	139	142
	$[AsnAla+H]^+$	126 (5)	132	139	134
	$[AsnVal+H]^+$	129 (5)	132	139	138
Dehydration	$[AsnGly+H]^+$	117 (6)	127	130	112
	$[AsnAla+H]^+$	105 (8)	119	122	98
	$[AsnVal+H]^+$	114 (5)	116	118	103

Table 5-5. *Comparison of 0 K Reaction Energies for the Deamidation and Dehydration of Asparagine Dipeptides. Calculations are performed using a 6-311+G(2d,2p) basis set with geometries calculated at the B3LYP/6-311+G(d,p) level. Energies are given relative to the ground structure of the reactant at each respective level of theory.*

5.4. Conclusion

The decomposition, and in particular the deamidation, of protonated AsnVal has been studied using a combination of infrared ion spectroscopy, guided ion beam tandem mass spectrometry, and quantum-chemical calculations. Loss of water and ammonia are the main fragmentation channels, which is consistent with results found for $[AsnGly+H]^+$ and $[AsnAla+H]^+$ (see ref. [59] and chapter 4). Molecular structures for the ionic decomposition products as well as energy dependent dissociation cross sections and transition state energies have been determined. We suggest that deamidation occurs via a nucleophilic attack of the backbone amide oxygen onto the amide carbon of the asparagine side chain, leading to the

formation of a reaction product incorporating a furanone moiety with a reaction energy of 129 ± 5 kJ/mol above the GS (within experimental uncertainty of predicted values at the B3LYP and MP2 levels). This furanone structure is in line with earlier studies on the deamidation of $[AsnAla+H]^+$ and $[AsnGly+H]^+$, although for those species a bifurcating mechanism was established, with the formation of a succinimide structure in parallel to the furanone product. Calculations of the rate limiting TSs for each of these species predict an increasing barrier towards succinimide formation with increasing alkyl side chain of the C-terminal residue: Val (145 – 172 kJ/mol) > Ala (137 – 163 kJ/mol) > Gly (133 – 151 kJ/mol), whereas the barriers towards the furanone product are independent of the C-terminal residue, qualitatively explaining the subtle differences in decomposition behavior for the three dipeptides. For $[AsnVal+H]^+$, the TS for succinimide formation is 7 – 40 kJ/mol higher than that for furanone formation.

Formation of furanone structures upon deamidation appears to be exclusive for gas-phase reactions. In condensed phase media, only formation of succinimide as an intermediate on the way to (iso-)aspartic acid has been reported. We speculate that the amide carbonyl is more strongly involved in H-bonding interactions in condensed media, reducing its nucleophilicity and thus reducing the likeliness of forming furanone structures. Further, on the basis of our structural analysis above, the relative energetics controlling deamidation appear to be related to the proximity of TS motions to the bulky Val isopropyl group. Thus, furanone formation (where TS motions are localized to the Asn side-chain) is energetically favorable compared to succinimide formation, where these motions are closer to the isopropyl group. Interestingly, for the succinimide reaction, the lowest energy TS orients the Val side-chain methyl groups towards either of the oxygens of the carboxylic acid (as observed in the $TS_{N\text{-}FUR}$), although one could imagine that such an orientation could sterically hinder the reaction as a result of the close proximity between the TS motions and the CH_3 group. However, rotation of the isopropyl group to other positions results in higher energy TSs. For instance, one TS was located where the isopropyl CH hydrogen was oriented towards the bond rupture/formation TS motions, although this TS was located ~10 kJ/mol higher in energy than $TS_{N\text{-}SUC}$.

Given these subtleties governing the deamidation reaction pathways of protonated AsnXxx peptides, it remains difficult to speculate on the outcome of the reaction for other Xxx residues (i.e., whether succinimide or furanone formation is observed). However, our current results do suggest that succinimide formation, the dominant pathway in solution, is controlled (at least in part) by steric effects of the

n+1 residue. Thus, reaction barriers for larger *n+1* residues are expected to increase, although our studies to date have only involved aliphatic side chains. One can imagine that polar side-chains having favorable or unfavorable inductive effects or that present more specific orientations resulting from hydrogen bonding may play a significant role in mediating the deamidation process. Investigations involving the *n+1* residues of Ser and Thr are currently underway in our laboratories.

Chapter 6

The deamidation reaction of asparagine-alanine and glutamine-alanine studied by chemical dynamics simulations

Chemical dynamics simulations have been employed to model the deamidation reactions of the protonated dipeptides Glutamine-Alanine (GlnAla) and Asparagine-Alanine (AsnAla). This reaction is, together with H_2O loss, a major fragmentation pathway in collision induced dissociation mass spectrometry (CID-MS) experiments on these systems. In a previous study using ion spectroscopy, we showed that while deamidation takes place from the side chain of the Asn and Gln residues, both peptides follow different reaction pathways (see chapter 4 / JASMS 2016). With the product ion structures firmly established experimentally, we can critically benchmark the results of the computational approach. For the deamidation of protonated GlnAla, a pyroglutamyl structure is formed by a nucleophilic attack of the N-terminal nitrogen, while for protonated AsnAla, a bifurcating mechanism was established, in which a nucleophilic attack by either the peptidyl amide oxygen generates a lactone structure, or a succinimide product ion is formed upon nucleophilic attack by the peptidyl amide nitrogen. In the simulations, the system is classically propagated on an “on the fly” computed potential energy surface (PES) using a semi-empirical Hamiltonian. The dynamics simulations are found to be capable of partially identifying the reaction pathways leading to the experimentally observed deamidated product structures, but some (mostly minor) dissociation channels remain unexplained.

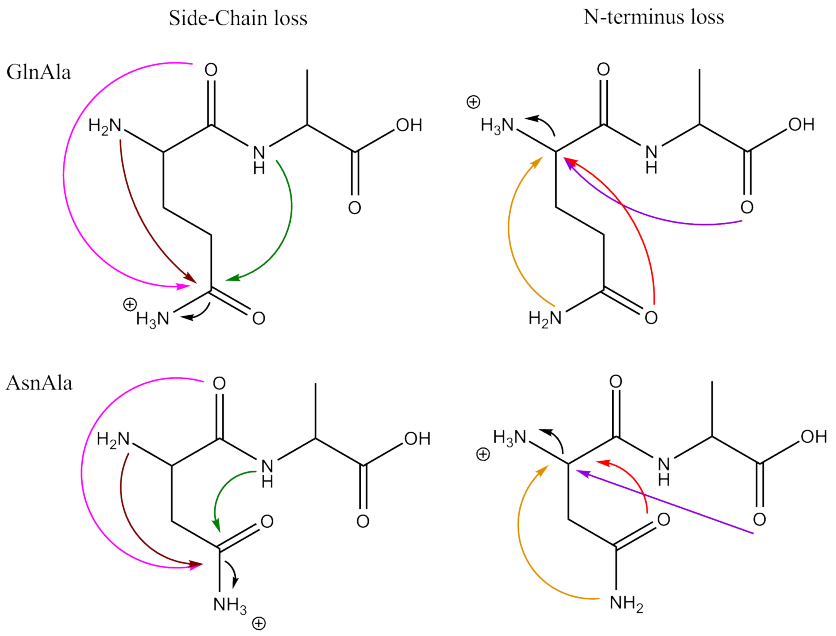
6.1. Introduction

On account of its high sensitivity and high resolving power, mass spectrometry (MS) has earned a role among analytical methodologies. Its main weakness, especially when compared with spectroscopic methods, lies in the limited chemical structure information contained within a mass spectrum. Apart from elemental composition information, the most common type of diagnostic information provided by (tandem) MS is a low-energy fragmentation mass spectrum. Although the fragmentation pattern of a molecule is likely a sensitive diagnostic for its structure, it can usually only be traced back to the molecular structure by use of a database of known fragmentation spectra. Unlike spectroscopic fingerprints (including NMR), MS/MS spectra are to date hardly reliably predictable from first principles. Recent work in various theoretical chemistry groups focusses on overcoming this challenge with the aim to accurately predict MS/MS spectra.

The fragmentation reactions of a molecular ion and the associated relative intensities of the product ions are contained within the full potential energy surface of the species. This surface can be probed using quantum-chemical calculations, where transition states can be determined and used in conjunction with statistical dissociation theories to predict the branching ratios between different dissociation channels and thus to predict an MS/MS fragmentation spectrum. However, even for relatively small systems such as the dipeptides considered here, quantum-chemical modeling of the dissociation reactions by mapping of the PES and locating the appropriate transition states is a complex and time-consuming task because of the many pre- and post-dissociation (conformational) rearrangements. In most theoretical studies modeling peptide dissociation reactions, only the most prominent transition states are investigated and only for some very small peptides has the entire reaction pathway been modelled [7,59,65,73,96,97]. Also, a good deal of "chemical intuition" is usually involved in locating the transition states.

Alternatively, the PES can be probed by Monte-Carlo type approaches. Our group has investigated the use of chemical dynamics simulations to characterize the reaction mechanisms of protonated peptides [66]. In these trajectory calculations, the atom positions are classically propagated on a potential energy surface that is calculated "on the fly" using a semi-empirical Hamiltonian [66,98-102]. Higher-level theoretical methods would require an excessive computational cost. The trajectories can lead to dissociated products depending on the energy provided to the system, and the resulting product ion structures can be analyzed. Here,

chemical dynamics simulations have been applied to study the deamidation reactions of two protonated dipeptides, as these systems are relatively small and their product ion structures have been spectroscopically determined (see chapter 4). This accurate knowledge of the product ion structures makes these systems an interesting testbed for trajectory simulations.



Scheme 6-1. *Potential mechanisms leading to the deamidation of protonated GlnAla (top) and AsnAla (bottom). Loss of NH_3 may occur from the N-terminus or from the side chain. Our previous study identified for the deamidation of protonated GlnAla that a nucleophilic attack of the N-terminal nitrogen on the Gln side chain carbonyl carbon takes place (brown arrow). For protonated AsnAla, the observed mechanisms are a nucleophilic attack of the peptidyl amide oxygen on the side chain (pink arrow) and a nucleophilic attack by the peptidyl amide nitrogen onto the side chain (green arrow).*

The low-energy collision-induced dissociation (CID) mechanisms of gaseous protonated peptides have been widely discussed in the literature [103,104] and are commonly described in terms of the mobile proton model [13]. Under low-energy collisions, the added proton can transfer from the most basic site in the peptide to other less basic sites, for example, from the N-terminal amine to a side chain moiety [6,18-20,57]. Hence, the deamidation reactions (NH_3 loss) for the dipeptides investigated here can conceptually take place from either the N-terminus or from the less basic side chain amide following proton migration. Scheme 6-1 presents

the possible reaction mechanisms for the deamidation of protonated GlnAla and AsnAla. Based on our previously published spectroscopic data ([63] or chapter 4), NH_3 is known to be expelled from the Gln or Asn side chain induced by a nucleophilic attack on the side chain amide carbon atom. Also, we derived that for protonated GlnAla, the attacking nucleophile is the N-terminal nitrogen atom leading to a pyroglutamyl product ion. In contrast, deamidation of protonated AsnAla proceeds along two bifurcating reaction pathways: 1) nucleophilic attack of the peptidyl amide oxygen on the side chain generates a structure featuring a 5-membered lactone ring, and 2) nucleophilic attack by the peptidyl amide nitrogen forms a succinimide ring containing product ion.

6.2. Computational methods

The molecular geometries of the N-terminally protonated precursor ions of GlnAla and AsnAla were optimized using density functional theory (DFT) at the B3LYP/6-31++G(d,p) level of theory in Gaussion09 revision D01 [31]. The potential energy surfaces were explored to identify the lowest-energy conformers using a Molecular Mechanics/Molecular Dynamics (MM/MD) approach employing AMBER 12 [32]. After minimization within AMBER, a simulated annealing procedure up to 500 K was used. Five hundred structures were obtained in this procedure and grouped based on structural similarity using appropriate rms criteria to give 20 to 30 candidate structures. Finally, these structures were each optimized at the B3LYP/6-31++G(d,p) level. Electronic energies were calculated at the MP2(full)/6-311+G(2d,2p)//B3LYP/6-31++G(d,p) level of theory at 298 K. A more detailed description of the procedure can be found elsewhere.[15]

6.2.1. Internal energy activation

Collisional activation of ions in MS/MS is accomplished through collisions with background inert gas molecules. The experimental set-up has been described in (chapter 2 and ref. [63]). In the ion-trap MS we assumed that the ion experiences a number of low-energy collisions that slowly and uniformly activate it prior to fragmentation. In order to simulate this process, we ran trajectory calculations where the ion has a specified amount of internal energy (see Ref. [98] for details) and we monitor the fragmentation products with the aim of identifying the fragmentation mechanisms. Note that in these trajectories the system is composed only of the energized ion itself, i.e., the collision itself is not modeled.

As starting geometries, we use the minimum-energy structures identified by the quantum-chemical calculations. For those structures, initial positions and momenta

were obtained by a classical microcanonical normal mode sampling [105] such that the activation energy was distributed among the ion's vibrational modes. We tested different amounts of activation energies for the precursor ion of AsnAla (225, 250, 275 and 300 kcal/mol). Trajectories were then propagated using the Velocity Verlet algorithm [106] to numerically solve Newton's equations of motion with a time step of 0.2 fs, which ensured energy conservation.

Energy (kcal/mol)	No reaction	Isomerization	Fragmentation	NH_3-loss
225	5 %	65 %	30 %	12 %
250	1 %	44 %	55 %	13 %
275	1 %	24 %	75 %	14 %
300	0 %	12 %	88 %	6 %

Table 6-1. *Percentages of observed isomerization, fragmentation and non-reactive trajectories during the chemical dynamics simulations for activation energies of 225, 250, 275 and 300 kcal/mol. Trajectories leading to NH_3-loss are given as a percentage of the fragmentation trajectories. The trajectory simulations were started with the proton positioned at the side chain amide group. As we are interested to follow NH_3 loss from the precursor ions, further chemical dynamics simulations have been employed by adding an activation energy of 275 kcal/mol.*

To compute energies and forces at each step of the trajectory, we employed the RM1 Hamiltonian[107], which was previously found to correctly describe fragmentation products and dynamics of amino acids and peptides [101,108-111]. An ensemble of 1000 trajectories was used for each system studied. These trajectories were propagated for up to 10 ps. The general chemical dynamics code VENUS[112], developed by Hase and coworkers, coupled with MOPAC[113] is used to perform the calculations. Trajectories were analyzed through an automatic approach[114,115] based on graph theory, which is able to discriminate whether at the end of the simulation a fragmentation or isomerization reaction has occurred, or if the ion retains the connectivity of the initial structure (i.e. no reaction). We classify the trajectory outcomes as either fragmentation, isomerization or no-reaction and Table 6-1 lists the percentage of occurrences of each of these for the different activation energies. The percentages shown in this table are results from simulations that started with the proton positioned at the side chain amide group as the 10 ps simulation time is not long enough to observe proton transfer and NH_3 loss. We observe that the number of reactive trajectories increases for higher activation energies and that a value of 275 kcal/mol gives the highest fraction of NH_3 loss events. This amount of energy has been used for all trajectory simulations.

6.3. Results and discussion

With the mobile proton model in mind, we may assume that dissociation occurs from structures with the proton in a position that is not the most basic site in the molecule. The computationally limited duration of the trajectories (10 ps) may not allow for proton migration and subsequent dissociation events. Trajectories are therefore started from initial structures with the added proton attached to different sites in the molecule. Since we are especially interested in deamidation reactions in this study, trajectories were initiated from structures having the proton either at the N-terminus (most basic site) or at the side chain nitrogen atom (corresponding to a higher-energy protomer of the precursor ion). Three types of trajectories are discerned during the chemical dynamics simulations: fragmentation of the precursor ion, proton transfer (isomerization) and non-reactive pathways, which are described in the following section.

Start	**Product**					
	N-term.	Side chain Nitrogen	Side chain Oxygen	Peptide Bond Nitrogen	Peptide Bond Oxygen	Other
GlnAla-N-term	X	4%	37%	3%	27%	29%
GlnAla-Side chain Nitrogen	69%	X	12%	6%	2%	11%
AsnAla-N-term	X	17%	43%	3%	20%	17%
AsnAla-Side chain Nitrogen	57%	X	22%	2%	11%	8%

Table 6-2. *Percentages of observed isomerization reactions for GlnAla and AsnAla with different initial protonation sites.*

6.3.1. Isomerization

The isomerization reactions for protonated GlnAla and AsnAla have been followed starting from structures with protonation at the N-terminus as well as with protonation at the side chain amide NH_2. Table 6-2 lists the most commonly occurring isomerization reactions and the percentage of trajectories that occurred

for each. The isomerization reaction that occurs most frequently is proton transfer from the N-terminus to the side chain oxygen. For both side chain protonated peptides, migration of the proton to the N-terminal nitrogen is the most frequently observed isomerization, which is not surprising since the N-terminus is the most basic site in the molecules.

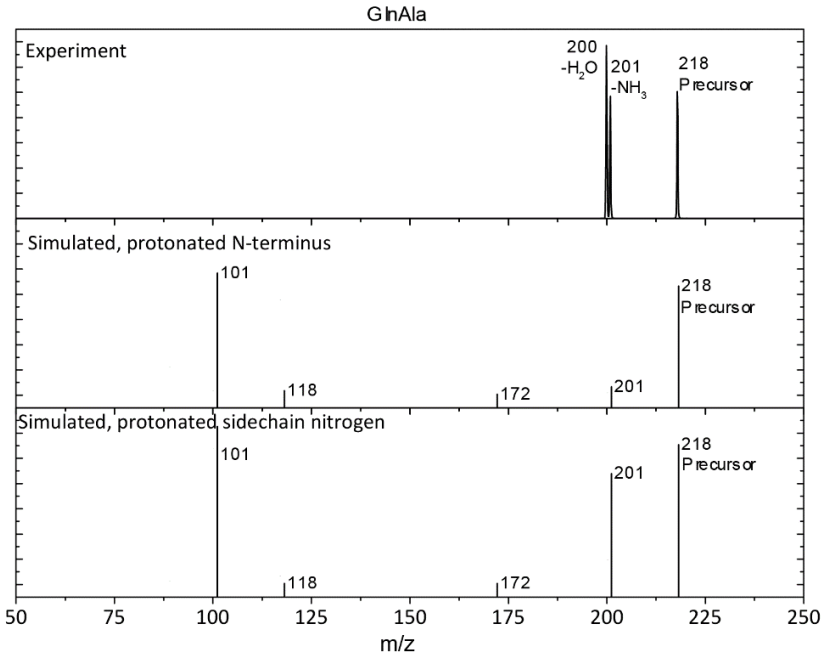
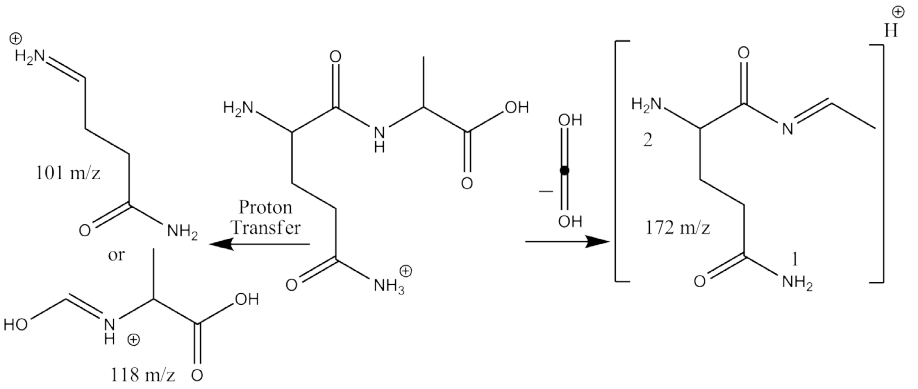


Figure 6-1. *The experimental CID mass spectrum for $[GlnAla+H]^+$ is shown in the top panel of this figure. The theoretical mass spectrum for GlnAla protonated at the N-terminus is presented in the middle panel of the figure, the theoretical mass spectrum for GlnAla protonated at the sidechain nitrogen of Gln is shown in the bottom panel of the figure.*

6.3.2. Fragmentation

The experimental CID mass spectra for protonated GlnAla and AsnAla are compared with the calculated CID spectra for the precursor ions protonated at the N-terminus and at the side chain nitrogen. Figure 6-1 presents the experimental (top panel) and simulated CID fragmentation patterns for GlnAla, with protonation at the N-terminus (middle panel) and the glutamine side chain (bottom panel). While the experimental spectrum only shows neutral loss of NH_3 and H_2O, the calculated MS spectra show additional peaks at m/z 172, (a_2-fragment), m/z 118 (x_2-fragment) and m/z 101 (a_1-fragment). Scheme 6-2 shows the structures of these

fragments. H_2O-loss is not observed in the simulations, while NH_3-loss is clearly more abundant starting from the side chain protonated precursor.



Scheme 6-2. *The structures of the fragments observed in the simulated mass spectrum of protonated GlnAla. The deamidated fragment is not specified here but in Scheme 6-2. During the simulations cleavage occurs without a cyclization reaction. Numbers indicate protonation sites in order of abundance.*

Figure 6-2 shows the fragmentation mass spectra for AsnAla obtained experimentally (top panel), by simulations starting from the N-terminally protonated structure (middle panel) and from the side-chain nitrogen protonated structure (bottom panel). In the experimental MS/MS spectrum, only NH_3 and H_2O neutral losses appear, while in the calculated spectra the y_1-fragment at m/z 89 and a_1-fragment at m/z 87 are the most abundant peaks. A full list of reaction products for the decomposition of AsnAla includes m/z 187 (-NH_3), m/z 159 (-NH_3-CO), m/z 158 (-H_2O-CO), m/z 145 (– NH_2 – CO – CH_3), m/z 118 (-NH-CH-CH_2-CO-NH_2), m/z 89 (-Gln), m/z 87 (-Ala- CO), m/z 72 (internal fragment) and m/z 70 (-Ala-CO-NH_3). These have the same structures as suggested by Boles et al.[97] for fragments at m/z 187, 158, 145 and 87. Fragments resulting from higher-energy pathways are not expected in our CID MS/MS experiments as the ions are subjected to slow-heating conditions by low-energy mass-resonant activation.

In TCID and SORI-CID the precursor ions may be subject to higher collision energies and different reaction products may thus be expected. Therefore, it appears that the TCID and SORI-CID mass spectra are more accurately predicted by our computational methods than the CID mass spectra studied here. In contradiction to the CID experiments, TCID makes use of a single-collision process and the added collision energy in TCID experiments can be accurately determined. The fragment

at m/z 115 has not been observed in the simulations, as well as sequential fragmentation products at m/z 124, 169, and 170. We suspect that this is due to the limited simulation time of 10 ps, which is too short to observe the proton transfer required to induce these fragmentation channels. Scheme 6-3 contains an overview of the structures of the fragments observed in the calculated spectra. In chapter 4 and 7, the structures of the fragment ions resulting from loss of NH_3 and H_2O from both GlnAla and AsnAla were spectroscopically elucidated.

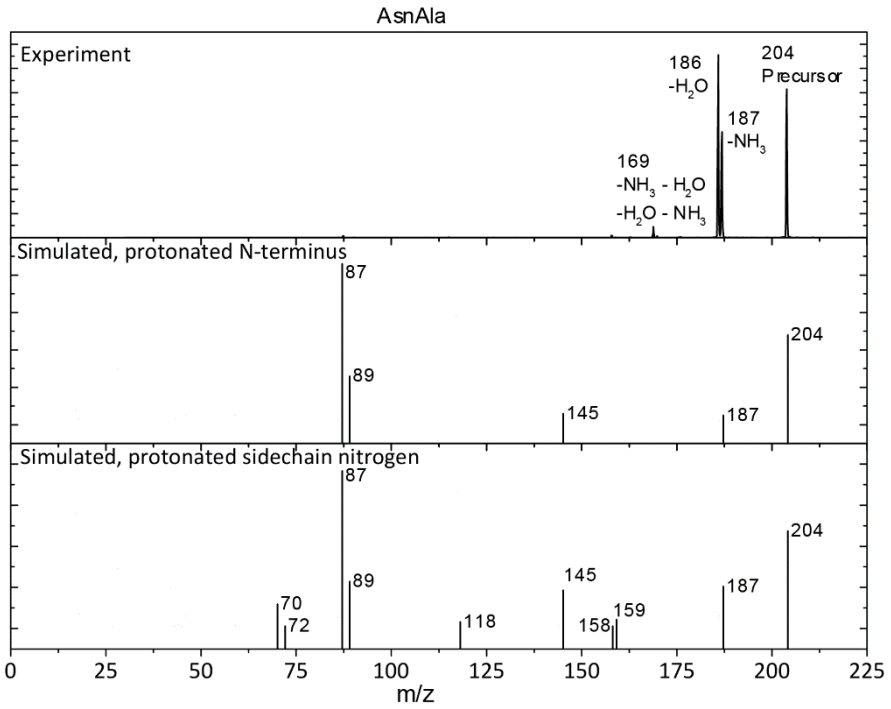
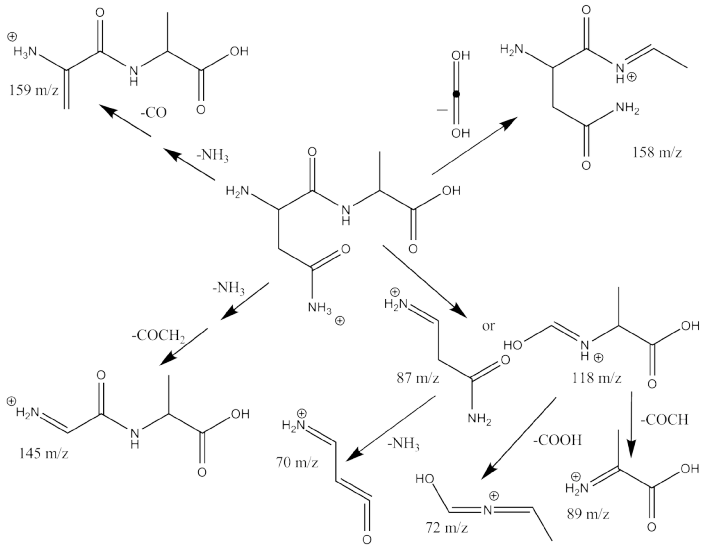


Figure 6-2. *The experimental CID mass spectrum for $[AsnAla+H]^+$ in the top panel is compared with theoretical mass spectrum for AsnAla protonated at the N-terminus (middle panel) and with the theoretical mass spectrum for AsnAla protonated at the side chain nitrogen (bottom panel).*

Because loss of H_2O requires transfer of the proton from the N-terminus to the side chain or to the C-terminus, loss of H_2O has not been observed for both dipeptides at the timescale of 10 ps used here. This also explains the higher intensity of the NH_3–loss mass peak in the simulated trajectories starting from the side chain nitrogen protonated system. Reversing the argument, this is a good indication that NH_3-loss occurs after the mobile proton has migrated to the side chain. The differences in the predicted fragmentation between initial structures with

protonation at the N-terminus and protonation at the side chain nitrogen are larger for AsnAla than for GlnAla. For AsnAla more fragments appear when the protonation is on the Asn side chain. This can possibly be explained by the shorter side chain and the therefore more efficient proton migration back to the backbone leading to the formation of some sequence ions (y_1, a_1, a_2 and x_2 ions) in the simulations.



Scheme 6-3. *The structures of the fragments observed in the simulated mass spectrum of AsnAla. The deamidated product structures are specified in Scheme 6-5. During the simulations cleavage occurs without a cyclization reaction.*

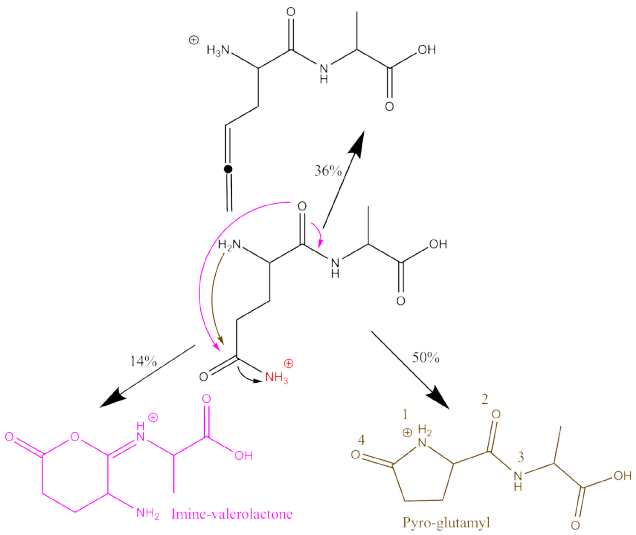
6.3.3. Deamidation mechanisms

We now focus on trajectories leading to deamidation starting from precursor ion structures with the proton positioned at the Gln or Asn side chain nitrogen atom.

Glutamine-Alanine (GlnAla)

The resulting product ion structures for the simulated deamidation reaction of protonated GlnAla are presented in Scheme 6-4. The NH_3 group is expelled from the Gln side chain independent of whether the precursor ion is protonated on the amide side chain or on the N-terminus. In the last case proton transfer to the side chain amide nitrogen is necessary for NH_3 loss to occur. The pyroglutamyl structure, formed by nucleophilic attack of the N-terminus on the side chain amide, is the most frequently formed product ion structure, in accordance with our previous ion spectroscopy study (chapter 4). However, simulations also show the

occurrence of an imino-valerolactone structure, formed by nucleophilic attack of the peptidyl oxygen on the side chain amide; this structure was explicitly not observed in the experiment (chapter 4). Direct cleavage of the side chain amide C-N bond, i.e. without nucleophilic attack, leads to the ketene structure, which occurs for a substantial fraction of the trajectories, but is also not observed experimentally. We suspect that running the trajectories longer would eventually convert this ketene structure into one of the cyclized product ions.



Scheme 6-4. *Reaction products resulting from the chemical dynamics simulations. The most frequently observed product ion is the pyroglutamyl structure, the imino-valerolactone structure has been observed in a minority of trajectories. The ammonia molecule is in all observed reactions lost from the glutamine side chain. Observed protonation sites are indicated by numbers in order of their presence.*

Table 6-3 lists the reaction enthalpy for the deamidation reactions and the excess in energy (added activation energy minus the reaction enthalpy) that is added during the simulations. The values in brackets present the reaction enthalpy at the MP2 level of theory. The excess of energy is 1013 kJ/mol for the pyro-glutamyl structure and 1014 kJ/mol for the imino-valerolactone structure at the RM1 level of theory. Compared to the MP2 level of theory, the energy differences between the structures listed in table 6-3 are relatively small for the semi-empirical calculations. These small energy differences between structures at the RM1 level of theory may explain the observation of structures in the dynamics simulations, that are not observed in the experiments and are not predicted at the MP2 level of theory.

	RM1	
Molecule	Reaction Enthalpy (kJ/mol)	Excess Energy (kJ/mol)
GlnAla-NH_3		
Linear	164 ([a])	987
Diketo-morpholine	156 (173)	995
Imino-dihydrofuran	183 (208)	968
Imino-valerolactone	137 (244)	1014
Glutamiride	140 (176)	1011
Pyro-glutamyl	138 (149)	1013
AsnAla-NH_3		
Linear	([b])	([b])
Diketo-morpholine	160 (120)	991
Furanone	177 (149)	974
Succinimide	103 (95)	1048

Table 6-3*. The amount of energy added to the molecules compared to the energy cost of product ion formation. The excess energy is defined as the CID energy (275 kcal/mol = 1150 kJ/mol) minus the reaction enthalpy. The reaction enthalpy at the MP2 level of theory is presented between the brackets (see chapter 4).* [a] *The reaction enthalpy of linear GlnAla-NH_3 structure could not be obtained at the MP2 level of theory due to the unstable nature of this structure.* [b] *The linear AnsAla-NH_3 structure undergoes a cyclization reaction to one of the other listed structures.*

For the deamidated products of GlnAla, the distances between the attacking nucleophiles (peptide bond oxygen/pink, N-terminal nitrogen/brown) and the side chain carbonyl carbon are plotted as a function of time in characteristic trajectories (see Figure 6-3). Additionally, the distances between the leaving NH_3-group and the carbonyl carbon atom is plotted with a black line. For the formation of the imino-valerolactone structure (top panel of Figure 6-3), loss of the NH_3 molecule occurs simultaneously with ring closure and no intermediate structure is observed. This reaction mechanism is concerted and synchronous. For the formation of the pyroglutamyl structure, an analogous concerted and synchronous reaction mechanism is observed, but in some trajectories an intermediate imino-valerolactone structure is observed prior to formation of the pyro-glutamyl structure. This reaction pathway is referred to as being synchronous but not-concerted and an example trajectory is shown in the bottom panel of Figure 6-3. Potentially, this could explain why the imino-valerolactone structure is not present in the experiment, as it rearranges to a pyroglutamyl structure according to the trajectory simulations, see bottom panel of Figure 6-3.

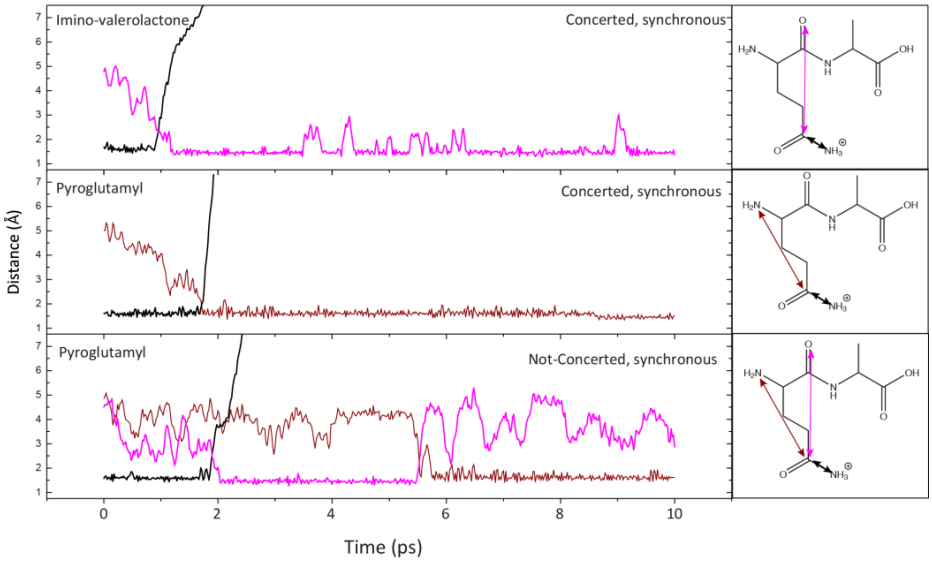
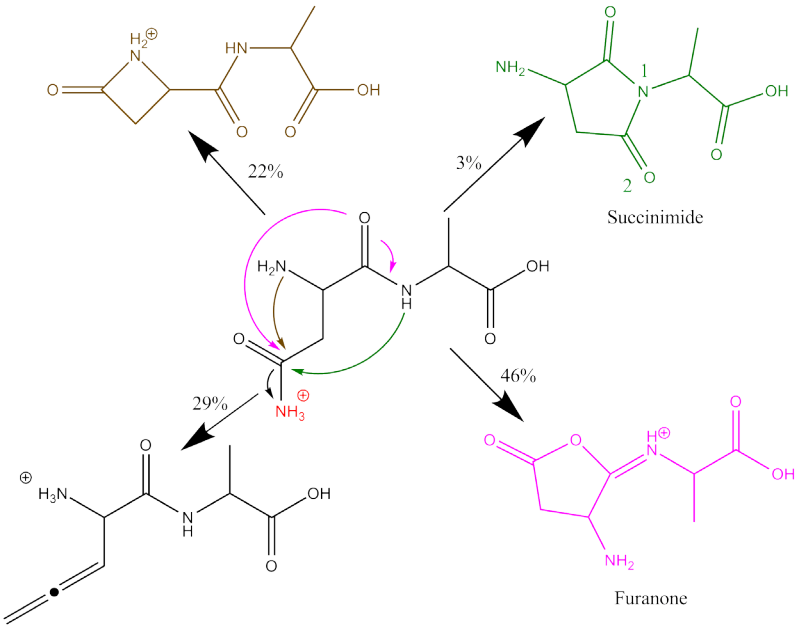


Figure 6-3. *The distance graphs for the formation of the imino-valerolactone and pyroglutamyl structures for three characteristic trajectories. The distance between the NH_3 leaving group and amide C-atom is plotted in black. The distance between the peptide bond oxygen and the side-chain amide carbon is shown in pink and that between the N-terminal nitrogen and the side-chain amide carbon in brown. The time is given in units of 0.02 ps.*

Asparagine-Alanine (AsnAla)

Scheme 6-3 presents the product ion structures resulting from deamidation of protonated AsnAla. Again independent of the initial protonation site, the NH_3 group is lost from the Asn side chain and the most frequently obtained product is the furanone structure. Although the succinimide structure is lower in energy by 46 kJ/mol, it was only formed in a few trajectories. Formation of both product ion structures agrees qualitatively with our experimental observations, but the succinimide-to-furanone ratio is approximately three times smaller in the simulations than in the CID experiment, as determined from the ratio of the unique MS^3 fragments for each of the structures. Transition State calculations showed that the rate limiting TS towards the succinimide structure is up to 32 kJ/mol higher than the TS leading to the furanone structure, explaining the dominant formation of furanone product ions.[97] The 4-membered ring structure in Scheme 6-5 is a stable structure during DFT optimization, although it is high in relative energy (+91 kJ/mol); it has not been observed experimentally. The computationally predicted ketene-containing linear structure, not involving cyclization, can again be rationalized by the short trajectory simulation time of 10 ps, insufficient for the full

reaction to complete. Furthermore, the linear structure is also formed after ring re-opening. Table 6-3 lists the excess energies at the RM1 level of theory, which is 974 kJ/mol for the furanone structure and 1048 kJ/mol for the succinimide structure.



Scheme 6-5. *The reaction products resulting from the Chemical Dynamics simulations. The most frequently observed product ion is the furanone structure, while the succinimide structure has been observed for a small number of trajectories, qualitatively in agreement with experimental observations. The ketene-containing structure is possibly an intermediate and has not been observed experimentally, nor has the 4-membered ring structure. In all observed reaction pathways the ammonia molecule is lost from the asparagine side chain. Observed protonation sites are indicated by numbers in order of their presence.*

Figure 6-4 shows selected interatomic distances during characteristic trajectories for the simulated deamidation of protonated AsnAla. The distance between the attacking nucleophile (N-terminal nitrogen/brown, amide bond oxygen/pink and amide bond nitrogen/green) and the side-chain carbonyl carbon are plotted. The distance between the leaving NH_3 group and the carbonyl carbon atom is represented by the black line in all panels. For the formation of a furanone structure, a concerted and synchronous reaction mechanism is observed in the bottom left panel and all other trajectories leading to this product ion structure. After loss of the NH_3 molecule, direct ring closure is observed. For the formation of

the 4-membered ring structure a concerted reaction mechanism has been observed that can be either synchronous (direct ring closure, top left panel) or asynchronous (NH_3 group leaves and ring closure happens after approximately 1 ps (left middle panel). The formation of a succinimide structure has been observed in only a small number (3) of trajectories, and the reaction mechanism is different in all three cases. In one case (top right) the reaction follows a concerted synchronous pathway, another case (middle right) shows a non-concerted synchronous reaction pathway, where formation of a succinimide proceeds via the higher-energy 4-membered ring structure. In the last case (bottom right) a non-concerted asynchronous pathway is observed, where the formation of the succinimide proceeds via the formation of a higher-energy furanone structure

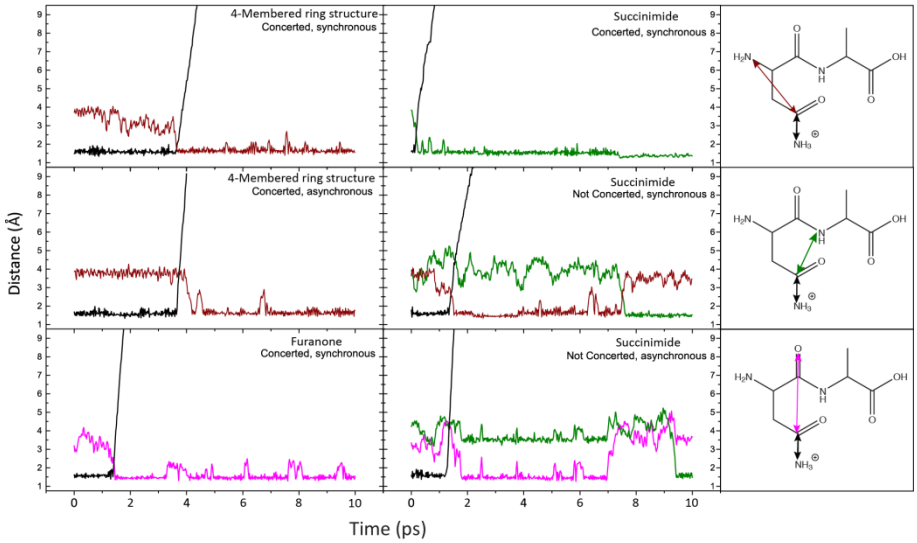


Figure 6-4. *Distance graphs for the formation of the 4-membered ring, furanone and succinimide product ions from protonated AsnAla. The distance between the NH_3 group and amide carbon from which it is detached in the reaction is plotted in black. The distance between the peptide bond oxygen and the Asn amide carbon is shown in pink, the distance between the N-terminal nitrogen atom and the Asn amide carbon in brown, and the green line represents the distance between the peptide bond nitrogen and the Asn amide carbon.*

6.4. Conclusion

According to the chemical dynamics simulations, deamidation of protonated GlnAla and AsnAla proceed along bifurcating reactions. Nucleophilic attack of the peptidyl amide oxygen on the side chain generates structures featuring a 5-membered furanone ring (AsnAla) or a 6-membered lactone ring (GlnAla). Alternatively, for AsnAla a succinimide product ion is formed upon nucleophilic attack by the peptidyl amide nitrogen onto the side chain carbonyl carbon, while for GlnAla, formation of a pyroglutamyl structure by attack of the N-terminus on the Gln amide carbon is predicted. Somewhat unexpectedly, the chemical dynamics also follow an analogous mechanism for AsnAla, leading to an awkward but stable 4-membered ring structure.

The formation of furanone, succinimide and pyroglutamyl structures for the deamidation of GlnAla and AsnAla are in accordance with pathways recently established by ion spectroscopy in chapter 4 and transition state calculations for deamidation of protonated AsnGly[59], AsnVal (chapter 5) and AsnThr. The 4-membered ring structure (AsnAla) and the 6-membered lactone structure (GlnAla) have not been observed in the experiments, however, these structures are among the least frequently observed structures in the chemical dynamics simulations.

Chapter 7

Dehydration reactions of protonated dipeptides containing asparagine or glutamine investigated by infrared ion spectroscopy

---

The role of specific amino acid side-chains in the fragmentation chemistry of gaseous protonated peptides resulting from collisional activation remains incompletely understood. For small peptides containing asparagine and glutamine, a dominant fragmentation channel induced by collisional activations is, in addition to deamidation, the loss of neutral water. Identifying the product ion structures from H_2O-loss from four protonated dipeptides containing Asn or Gln using infrared ion spectroscopy, mechanistic details of the dissociation reactions are revealed. Several sequential dissociation reactions have also been investigated and provide additional insights into the fragmentation chemistry. While water loss can in principle occur from the C-terminus, the side chain or the amide bond carbonyl oxygen, in most cases the C-terminus was found to detach H_2O, leading to a b_2-sequence ion with an oxazolone structure for AlaGln, and bifurcating mechanisms leading to both oxazolone and diketopiperazine species for AlaAsn and AsnAla. In contrast, GlnAla expels water from the amide side chain leading to an imino-substituted prolinyl structure.[*]

[*] Adapted from: Kempkes, L. J. M.; Martens, J.; Berden, G.; Oomens, J., Dehydration reactions of protonated dipeptides containing asparagine or glutamine investigated by infrared ion spectroscopy. *Int. J. Mass Spectrom.* **2018,** *429*, 90-100.

7.1. Introduction

Tandem mass spectrometry has become an indispensable tool in proteomics as it enables the routine determination of the amino acid sequence of peptides and proteins. This process involves the gas-phase dissociation of the protonated peptide and an analysis of the m/z values of the resulting fragment ions. In its most common form, dissociation is accomplished through collisional activation. Collision-induced dissociation (CID) primarily leads to backbone dissociation at the amide bonds, generating sequence ions referred to as b-type (N-terminal) or y-type (C-terminal) [11-13,104,116-119]. The molecular structures of the resulting b-type ions have been under particularly vigorous debate and it has now been established that b-type ions most often possess an oxazolone structure; for b_2-fragment ions, a diketopiperazine structure is in some cases formed alternatively [120]. Collision-induced migration of a proton from the most basic site to the amide bond where dissociation occurs, accompanied by a nucleophilic attack by the adjacent backbone carbonyl oxygen or the N-terminal nitrogen are suggested as the reaction pathways to these product ion structures, respectively [18-20]. Deviations from this general behavior have also been observed, for instance, b-ions from GlyGlyGly contain both five- and eight-membered ring structures [99]. Furthermore, the side chain can become actively involved in the reaction mechanism for instance in the formation of b_2 ions from protonated PheGlnAla (a glutarimide structure) and AlaAsnAla (a succinimide structure) [11].

Small molecule loss upon CID often occurs in addition to backbone fragmentation, where ammonia (deamidation) and water (dehydration) loss are the most common examples. The resulting non-sequence ions are typically not used for protein identification as their m/z values are often not recognized by sequencing algorithms [19,54,64]. However, since these processes are so common, a mechanistic understanding is of importance and may aid in further improving the interpretation of CID MS/MS spectra and the accuracy of sequencing [116].

In chapter 4, we investigated the deamidation of four very similar dipeptides in order to compare the influence of the side chain identity – Asparagine (Asn) versus Glutamine (Gln) − and the amino acid sequence − AlaXxx versus XxxAla, where Xxx is Asn or Gln. Here, we complement these studies by an investigation of the elimination of water from these same four peptides, so that we are able to completely map out their dissociation chemistry. Structures of the product ions were determined using a combination of mass spectrometry, infrared multiple

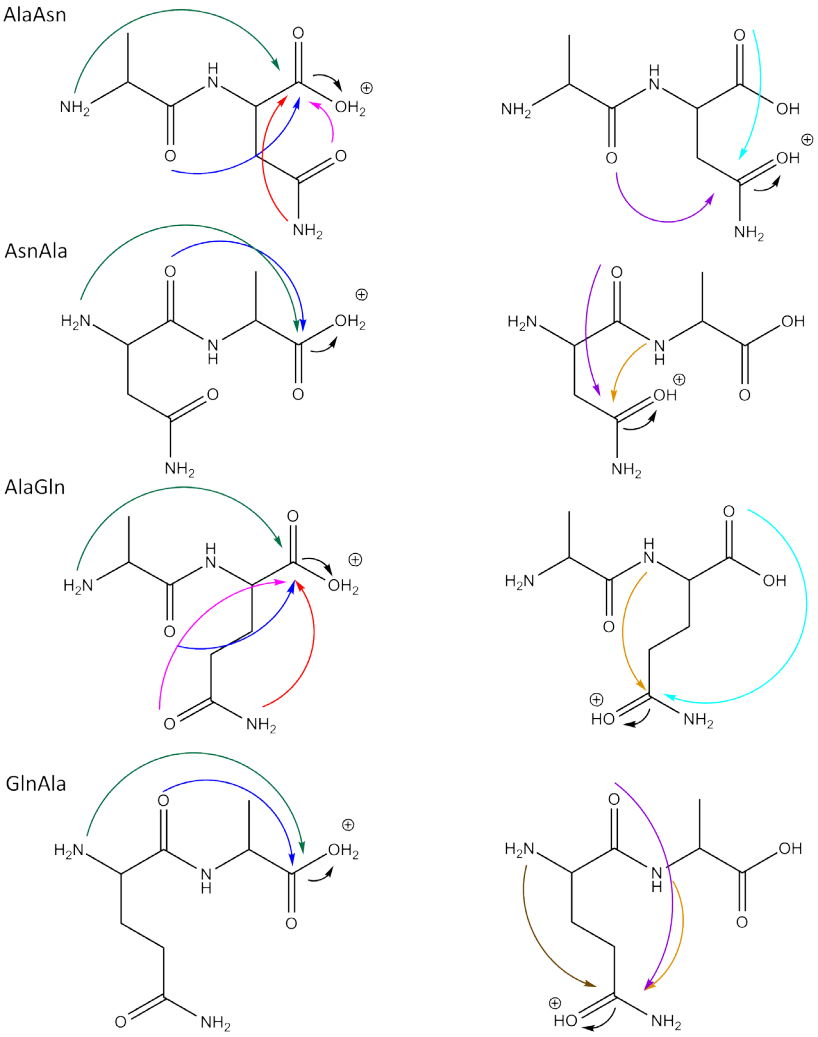
photon dissociation (IRMPD) spectroscopy and theoretical calculations. IRMPD spectroscopy has previously been used to determine the structures of several protonated peptide precursor ions and fragments [45,49,52,121-123], and has proven to be a powerful method to distinguish isomers, (prototropic) tautomers and conformers [23-25,30,38,44-46,48-50,53,54].

Various studies have examined the product structures of water loss from different protonated peptides[124-126], considering in particular the influence of the side chain [11,12,120], peptide length and the identity of the terminal residue in the departing fragment [120]. Note that water loss from an *n*-residue peptide gives a product ion corresponding to the b_n sequence ion (at least in terms of its m/z value). Water loss from ArgGly and AsnGly has been shown to lead to an oxazolone b_2 fragment, while for GlyArg it leads to a diketopiperazine structure [59,127]. The b_2 ions of HisAla bifurcate to both oxazolone and diketopiperazine fragment ions [47,120]. However, for protonated GlyGlyGlyGly, water loss occurs from the N-terminal amide group, initiated by a nucleophilic attack of the neighboring amide nitrogen, and hence does not result in an oxazolone structure [128]. While, GlyGly dissociation does not form a b-ion, b-ions from GlyGlyGly contain both five- and eight-membered ring structures [99]. Tripeptides that contain Asn or Gln have been shown to dissociate along different fragmentation mechanisms [11,12,120].

In order to understand why the formation of b_2-ions from AsnAlaAla and AlaAsnAla follow different reaction mechanisms, we have selected simplified dipeptide systems. These insights are valuable for understanding the fragmentation behavior of small peptides, and can be extrapolated to explain the fragmentation mechanisms of larger peptides and proteins and can as well be used to predict which reaction mechanisms lead to the formation of b_2 ions and which lead to the loss of small neutrals.

For the four dipeptides investigated here, Alanine-Asparagine (AlaAsn), Asparagine-Alanine (AsnAla), Alanine-Glutamine (AlaGln) and Glutamine-Alanine (GlnAla), loss of neutral ammonia and water are the dominant reaction channels upon collisional activation. The dehydration reaction of these dipeptides can lead to b_2-type ions if loss of water occurs from the C-terminus, however, water loss can potentially also occur from the peptide bond carbonyl or from the side-chain. Dehydration via the amide side chain can take place following a reaction that is similar to the reaction mechanism for the loss of ammonia (see chapter 4). To support the identification

of the dehydrated ion structures, secondary fragmentation reactions were also studied providing further confirmation for the identified structures and allowing us to construct a comprehensive map of the reaction network for each of the peptides is presented.



Scheme 7-1. *Possible nucleophilic attack rearrangements leading to dehydration. Both the attack on the C-terminus (left column) as well as on the side chain (right column) are considered. Mechanisms leading to energetically less favorable 4-,7- or higher membered ring structures are not shown here as they were consistently found to lead to high-energy fragment structures. Analogous reaction mechanisms are indicated by arrows of the same color.*

7.2. Experimental and computational methods

7.2.1. IRMPD spectroscopy

To obtain infrared multiple-photon dissociation (IRMPD) spectra of the dehydrated peptide fragment ions, a modified 3D quadrupole ion trap mass spectrometer (Bruker, AmaZon Speed ETD) was used [60] in combination with the Free Electron Laser for Infrared eXperiments (FELIX) [27]. IRMPD spectra were recorded over the 800-2000 cm^{-1} region. For IRMPD spectra in the hydrogen stretching range (3200-3800 cm^{-1}), an optical parametric oscillator (OPO, Laser Vision, Bellevue, USA) was used [15]. Employing FELIX, ions were irradiated with 6 μs long macropulses at a repetition rate of 10 Hz, each pulse having an energy of approximately 40 mJ and a bandwidth of ~0.5% of the center frequency. The Nd:YAG-pumped OPO generates 5 ns long pulses of approximately 15 mJ at 10 Hz and has a bandwidth of approximately 3 cm^{-1}. The frequency of FELIX is calibrated using a grating spectrometer and a wavemeter is used for the OPO. IRMPD spectra are corrected for fluctuations of the pulse energy over the scan range assuming a linear power dependence of the fragmentation signal.

Protonated peptide ions were generated using electrospray ionization (ESI) from 10^{-6} M solutions in 50:50 acetonitrile:water with ~0.1% formic acid added. Collisional activation of the isolated dipeptides for 40 ms with an amplitude parameter of approximately 0.3 V generated the $[M+H-H_2O]^+$ ions, which were subsequently mass selected and irradiated with two (more for OPO measurements) pulses from the IR laser. The IR induced fragment yield at each wavelength is obtained from five averaged mass spectra and calculated by relating the parent and fragment ion intensities according to $\sum I(fragments)/\sum I(fragments + parent)$ [26,60]. Peptides were purchased from GeneCust (Luxembourg) and used without any further purification.

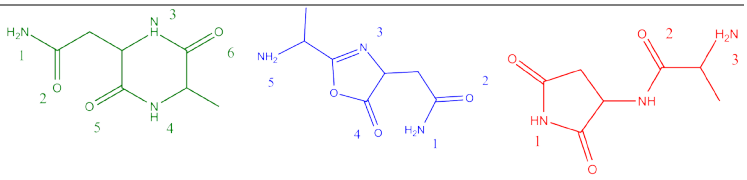
7.2.2 Computational chemistry

Density Functional Theory (DFT) calculations were performed at the B3LYP/6-31++G(d,p) level of theory using Gaussian 09 revision D01 [31]. Molecular geometries of the peptide fragment ions were optimized for different conceivable protonation sites and their linear IR spectra were predicted. Computed harmonic vibrational frequencies were scaled by 0.975 and convoluted with a 25 cm^{-1} full-width-at-half-maximum (FWHM) Gaussian line shape to facilitate comparison with experimental spectra. In order to obtain the lowest energy conformers and to explore the

potential energy surface, a molecular mechanics/molecular dynamics (MM/MD) approach using AMBER 12 [15,32,38,60] was applied on each peptide fragment ion. MS^3 fragments were considered to be small enough that no MM/MD approach was required. After an initial MM geometry optimization within AMBER, a simulated annealing procedure up to 500 K was used to obtain 500 structures. These structures were grouped based on their structural similarity using appropriate rms criteria to give 20-30 candidate structures. These structures were then optimized using the DFT protocol and their scaled harmonic spectra were compared with the experimental spectra. The computational procedure is described in more detail elsewhere [15,38,60].

7.3. Results and discussion

7.3.1. Dehydrated Alanine-Asparagine, $[AlaAsn+H - H_2O]^+$



Name → ↓ $[H]^+$ Site	1 Diketopiperazine	2 Oxazolone	4 Succinimide
1	[a]	+ 185	+ 186
2	+ 59	+ 157	+ 35
3	+ 66	+ 50	+ 20
4	+ 103	+ 143	
5	**0**	+ 59	
6	+ 29		

Table 7-1. *Calculated relative Gibbs free energies in kJ/mol for the different possible isomeric structures resulting from H_2O loss from protonated AlaAsn, each given for different conceivable protonation sites. The colors of the structures correspond with the colors of the arrows in Scheme 7-1.*

[a] *Proton migrates to position 5 during optimization.*

As depicted in Scheme 7-1, several possible mechanisms leading to the expulsion of H_2O were considered. Table 7-1 presents the three lowest energy structures arising from these calculations. Additional structures with their relative energies can be found in Table S1 in ref. [129]. The colors of the molecules in the tables correspond to the colors of the arrows indicating the nucleophilic attacks in Scheme 7-1. For each of the resulting structures, multiple protonation sites are conceivable as indicated by numbers 1 - 6 in the first column of Table 7-1. Structure **1.5**, the oxygen-protonated diketopiperazine structure, is the overall lowest energy structure.

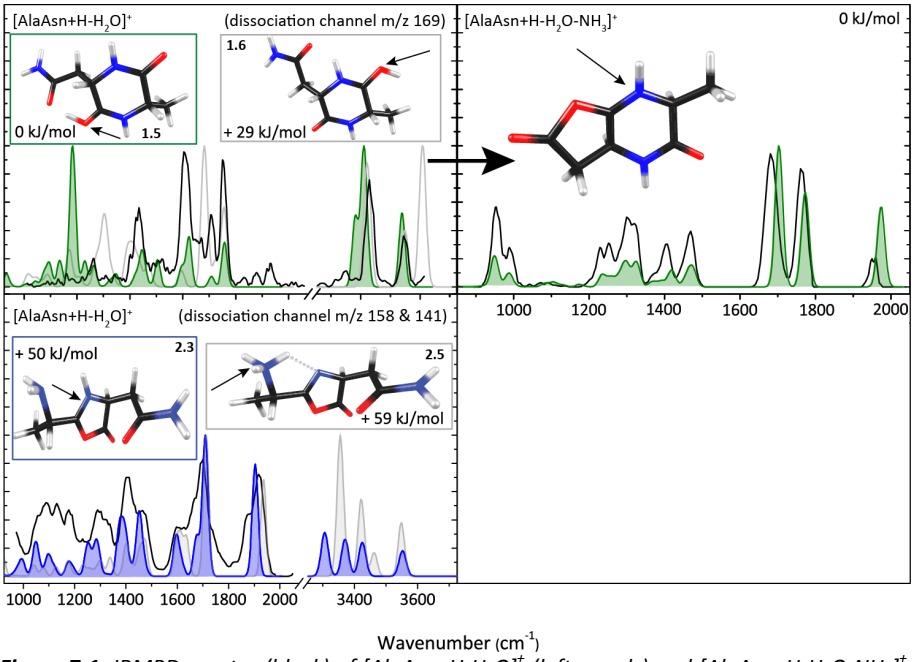
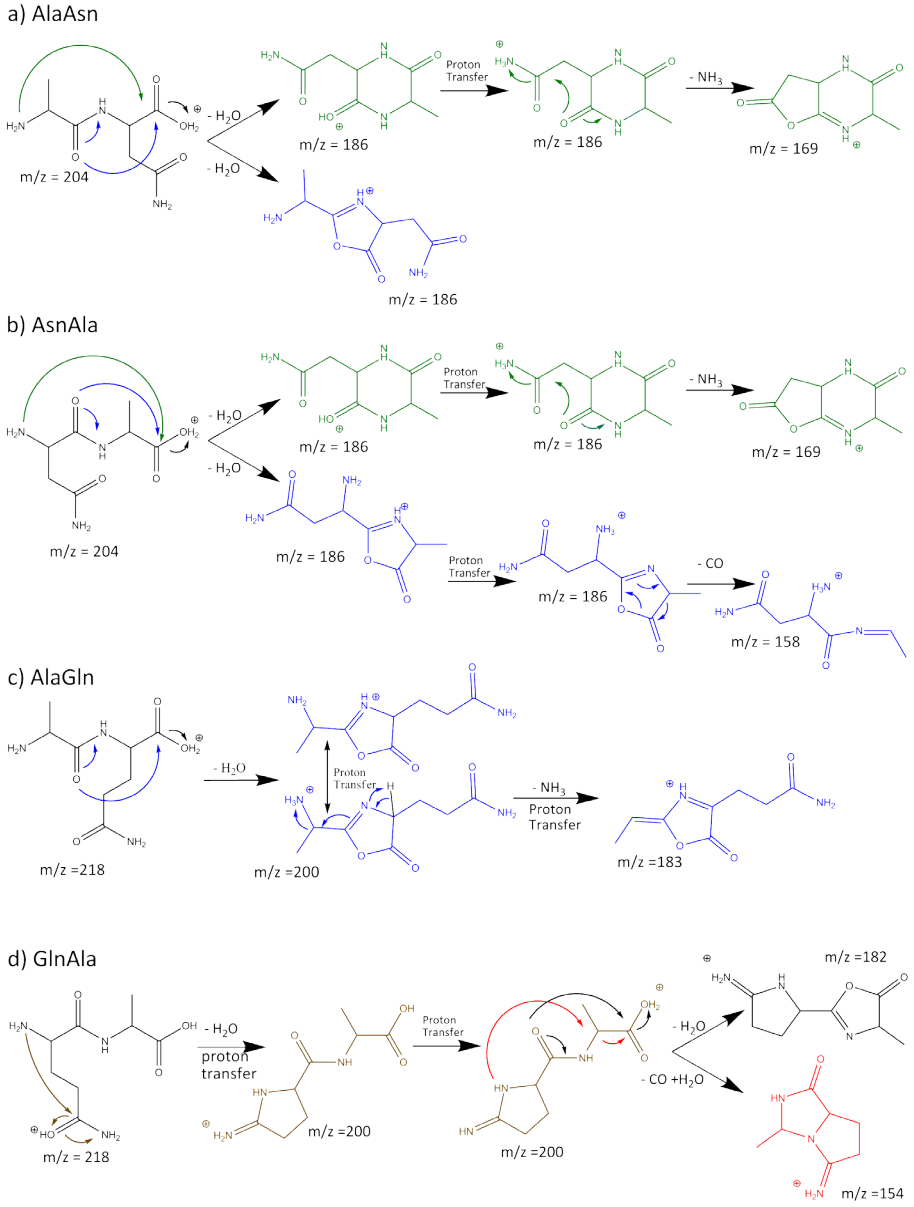


Figure 7-1. *IRMPD spectra (black) of $[AlaAsn+H-H_2O]^+$ (left panels) and $[AlaAsn+H-H_2O-NH_3]^+$ (right panel) compared to calculated spectra. The IRMPD spectra of $[AlaAsn+H-H_2O]^+$ in the left panels are generated by plotting the wavelength-dependent fragmentation into the m/z 169 channel (top, loss of NH_3) and that into the m/z 158 and 141 channels summed (bottom, loss of CO and CO+NH_3). Experimental spectra are plotted along with the best matching theoretical spectrum, where colors correspond to Table 7-1. The spectrum of the $[AlaAsn+H-H_2O-NH_3]^+$ MS^3 fragment ion (top right) confirms its structure as being formed from the diketopiperazine. Protonation sites are indicated with arrows.*

For $[AlaAsn+H–H_2O]^+$, two distinct fragments are produced by IRMPD, one at m/z 169 (corresponding to loss of NH_3, top left panel in Figure 7-1) and the other at m/z 158 (loss of CO, bottom left panel). The latter is accompanied by a minor dissociation into m/z 141 corresponding to additional NH_3 loss. Determining the IRMPD yield separately for these two channels gives two distinct IRMPD spectra, suggesting that multiple isomers are present in the population of $[AlaAsn+H–H_2O]^+$, each following distinct fragmentation pathways.

The band near 1925 cm^{-1} in the spectrum encoded into the CO and $CO+NH_3$ loss channels (Figure 7-1, bottom left) is the characteristic C=O stretch of an oxazolone ring thus identifying the oxazolone structure **2**. In fact, this band shows a splitting which suggests that the two lowest-energy oxazolone tautomers co-exist and the spectrum between 1000 and 1800 cm^{-1} agrees with the calculated IR spectra of both **2.3** and **2.5**. The relative Gibbs free energy difference between them is only 9 kJ/mol, which makes transfer of the proton from the oxazolone ring to the N-terminus thermodynamically possible as has been previously observed [49,120]. The band at 1710 cm^{-1} is assigned to C=O stretching of the side chain amide and the band around 1600 cm^{-1} is assigned to NH_2 bending at the N-terminus.

In contrast, the absence of the oxazolone band in the spectrum around 1900 cm^{-1} encoded into the NH_3-loss channel is a strong indication that the carrier has a diketopiperazine structure (**1**). The small peaks in the 1800-2000 cm^{-1} range are likely due to minor NH_3 loss from the oxazolone structure. Protonation on each of the distinguishable diketopiperazine oxygens leads to different H-bond stabilized structures (**1.5** and **1.6**) and comparison with their computed spectra in green and grey suggests the presence of more than one tautomer is evident [120]. The band around 1760 cm^{-1} in the calculated spectra of both tautomers is due to the unprotonated carbonyl C=O stretch. The band near 1420 cm^{-1} is due to CH bending in the CH_3-group. Several weaker bands predicted in the 1000 - 1400 cm^{-1} range are barely visible in the experimental spectrum. The intense band predicted at 1180 cm^{-1} is not observed in the experiment; the computation shows that it is due to a normal mode with dominant $O–H^+\cdots O$ bending character, which we suspect to be subject to significant spectral broadening and shifting in the IRMPD spectrum and to poor modelling by the harmonic calculations [130-135]. Other isomers/tautomers have been considered, but no convincing spectral match was found (see Figure S1 in in ref. [129]) and these structures are not further discussed here.



Scheme 7-2. *Proposed reaction mechanisms for H_2O-loss from the dipeptides investigated as derived from the product ion structures identified based on their IRMPD spectra.*

The right panel of Figure 7-1 shows the IRMPD spectrum of the MS^3 fragment $[AlaAsn+H-H_2O-NH_3]^+$ at m/z 169. Based on the agreement of the experimental spectrum with the predicted spectrum for a structure arising from NH_3 loss from the amide side chain of **1**, we assign the structure of this secondary fragment as a lactone/diketopiperazine fused ring structure. The band around 1975 cm^{-1} is due to carbonyl C=O stretching in the lactone ring, the 1775 cm^{-1} band is assigned to C=O stretching in the diketopiperazine ring, and the band near 1710 cm^{-1} is CN stretching of the protonated nitrogen. This assignment lends further support to the identification of $[AlaAsn+H–H_2O]^+$ as **1**. The intensity of the $[AlaAsn+H–H_2O-CO]^+$ product was too low to investigate its structure by IRMPD spectroscopy.

Scheme 7-2a shows the bifurcating reaction mechanism in the dehydration of protonated AlaAsn suggested by the present spectroscopic observations. Both isomers are formed through H_2O-loss from the C-terminus. To form the diketopiperazine isomer, nucleophilic attack occurs from the N-terminus, while the oxazolone isomer is formed via nucleophilic attack from the peptide bond oxygen. These pathways thus follow the well-known *b*-*y* fragmentation mechanism for protonated peptides [13]. Subsequent MS^3 dissociation of the diketopiperazine expels NH_3 from the Asn side chain. The oxazolone isomer expels CO, probably leading to an *a*-type sequence ion, although no spectroscopic information is available here (see further in the discussion of AsnAla).

Assuming that the diketopiperazine structure dissociates exclusively via NH_3 loss (into m/z 169) and that the oxazolone dissociates exclusively via loss of CO and $CO+NH_3$ (into m/z 158 and 141), we can use the CID MS^3 spectrum to estimate the relative dissociation rates into the bifurcating oxazolone and diketopiperazine pathways upon dehydration of the precursor peptide $[AlaAsn + H]^+$. Figure 7-2 shows that this ratio is about 4.5 to 1 in favor of the diketopiperazine pathway.

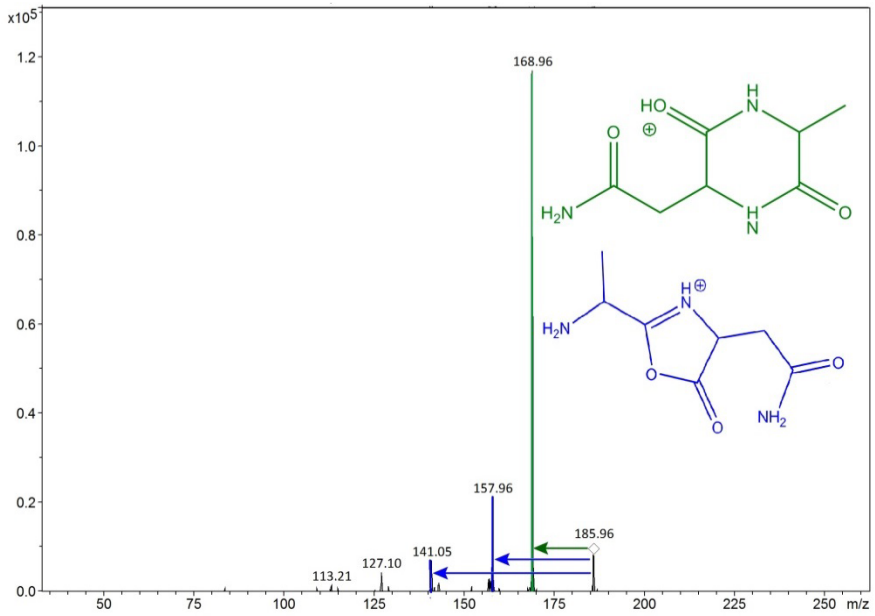
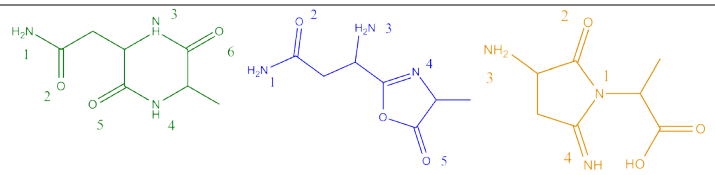


Figure 7-2. *MS^3 CID mass spectrum of $[AlaAsn + H - H_2O]^+$ at m/z 186. From the IRMPD spectra, we know that the diketopiperazine isomer (**1**, green) undergoes NH_3-loss to form m/z 169, while the oxazolone isomer (**2**, blue) detaches CO and CO+ NH_3 at m/z 158 and m/z 141. This MS^3 spectrum then suggests an approximately 4.5-to-1 ratio of the diketopiperazine and oxazolone isomer of the ion at m/z 186.*

7.3.2. Dehydrated Asparagine-Alanine, $[AsnAla + H - H_2O]^+$

For $[AsnAla+H-H_2O]^+$, Scheme 7-1 presents four possible reaction mechanisms leading to isomers **6-9**, for each of which all potential protonation sites have been considered; computed relative Gibbs free energies are shown in Table 7-2 for the most prominent ones, with additional structures given in Table S2 in ref. [129]. Ions **6** and **7** arise from H_2O loss from the C-terminus, while **8** and **9** form upon H_2O loss from the side chain. Figure 7-3 shows the IRMPD spectrum of the dehydrated product ion $[AsnAla + H - H_2O]^+$ at m/z 186, with the IR dissociation into channel m/z 169 (loss of NH_3) and that into m/z 158 + m/z 141 (loss of CO and CO + NH_3) again plotted separately.



Name →	6	7	9
↓ $[H]^+$ Site	Diketopiperazine	Oxazolone	Iminopyrrolidinone
1	+ 72	+ 198	+ 571
2	+ 20	+ 164	+ 146
3	[a]	+ 58	+ 79
4	+ 115	+ 43[b]	+ 13
5	**0**	+ 198	
6	+ 48		

Table 7-2. *Calculated relative Gibbs free energies in kJ/mol for isomeric structures of the H_2O-loss product from protonated AsnAla for different protonation sites*

[a] Proton migrates to position 2 during optimization.

[b] Multiple conformers identified; lowest energy is listed.

The spectra encoded into each of these mass channels are substantially different, suggesting again two different m/z 186 precursor ion structures to co-exist. The spectrum observed in channel m/z 169 (top left in Figure 7-3) is attributed to diketopiperazine structure **6**. Note that head-to-tail cyclization of AlaAsn and AsnAla upon H_2O-loss lead to the same structure, so that **6** is in fact equal to **1**. The experimental spectrum matches reasonably well with the spectrum computed for tautomer **6.5**, with the sharp features in 1400-1800 cm^{-1} range reproduced well and the 1000 - 1400 cm^{-1} range characterized by a series of weaker and unresolved bands. The strong band in the computed spectrum near 1180 cm^{-1} is not observed in the experiment as it is due to a normal mode with substantial $O–H^+\cdots O$ bending character, as discussed above. Our structure assignment is thus the same as that for the $[AlaAsn+H–H_2O]^+$ fragment ion presented in Figure 7-1, top left panel. Although the sharp, strong bands in the two experimental spectra coincide closely, discrepancies between the two experimental spectra are also observed. We

attribute these discrepancies to the different laser pulse energies used in both experiments, with that for the $[AlaAsn+H–H_2O]^+$ fragment being approximately three times lower, suppressing the weaker bands in the spectrum. The IRMPD spectrum of the MS^3 fragment ion $[AsnAla+H–H_2O–NH_3]^+$ in the top right panel of Figure 7-3 suggests that its structure is indeed identical to that established for $[AlaAsn+H–H_2O–NH_3]^+$ and thus provides further support for the structural identification of the MS^2 ion.

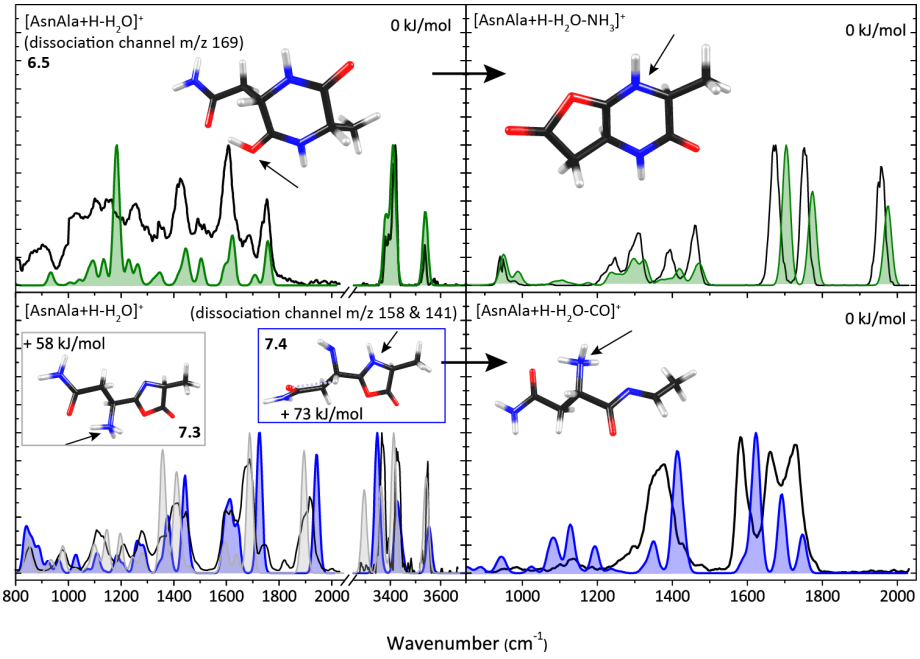


Figure 7-3. *IRMPD spectra (black) of $[AsnAla+H–H_2O]^+$ (left), $[AsnAla+H–H_2O–NH_3]^+$ (top right) and $[AsnAla+H–H_2O–CO]^+$ (bottom right) compared with the closest matching calculated spectra. The spectrum at the top left is generated from IR dissociation channel m/z 169 (sequential loss of NH_3), that at the bottom left is derived from reaction channels m/z 158 and 141 (sequential loss of CO and CO+NH_3). Protonation sites are indicated with an arrow.*

The IRMPD spectrum derived from IR induced dissociation channels m/z 158 and 141 (Figure 7-3, bottom left) possesses a strong and diagnostic feature centered around 1900 cm^{-1}, suggesting an oxazolone structure. As was seen for $[AlaAsn+H-H_2O]^+$ [49,120], this feature is broadened and likely consists of two unresolved bands due to two oxazolone tautomers, **7.3** and **7.4,** whose calculated spectra are

reproduced in grey and blue. For **7.4** (+73 kJ/mol), an alternative conformer substantially lower in energy (+43 kJ/mol) was identified and its computed spectrum is shown in Figure S2 in ref. [129]. The two spectra differ primarily around 1700 cm^{-1}, but the experimental spectrum does not allow us assign either one or the other conformer, especially when taking into account the additional contribution from **7.3**, which also has a strong band in this range.

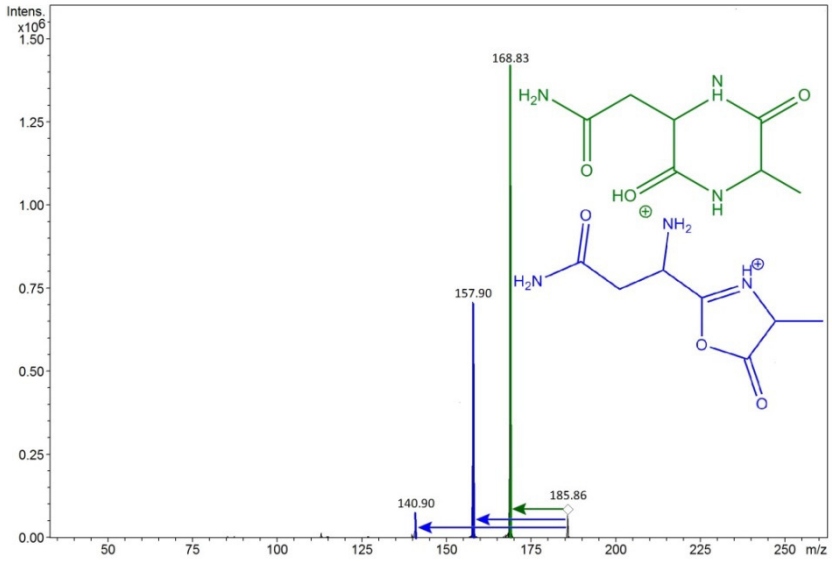


Figure 7-4. *MS^3 CID mass spectrum of $[AsnAla + H - H_2O]^+$ at m/z 186. From the IRMPD spectra, we know that the diketopiperazine isomer (**6**, green) undergoes NH_3-loss to form m/z 169, while the oxazolone isomer (**7**, blue) detaches CO and CO+ NH_3 at m/z 158 and m/z 141. This MS^3 spectrum then suggests an approximately 2-to-1 ratio of the diketopiperazine and oxazolone isomer of the ion at m/z 186.*

The oxazolone fragment undergoes MS^3 dissociation by sequential CO-loss, as is common for oxazolone *b*-type sequence ions [13]. The spectrum of this MS^3 fragment ion (bottom right in Figure 7-3) exhibits bands at 1747 cm^{-1}, C=O stretching of the peptide linkage, 1691 cm^{-1}, C=O stretching of the side chain amide, 1628 cm^{-1}, imine C=N stretching, and 1619 cm^{-1}, NH bending in the NH_3 group. The ion thus has the typical imine structure of *a*-type sequence ions, though with the proton on the N-terminus [136].

Scheme 7-2b summarizes the suggested reaction mechanism for the loss of H_2O from protonated AsnAla and the sequential losses of NH_3 and CO. Dehydration of

protonated AsnAla occurs on the C-terminus and follows bifurcating pathways analogous to those identified for protonated AlaAsn, leading to oxazolone and diketopiperazine isomers. Further activation of the diketopiperazine expels NH_3 from the Asn side chain, forming a 5/6 membered bicyclic structure. In contrast, the oxazolone eliminates CO forming a linear *a*-type sequence ion incorporating an imine group. Inspection of the CID MS^3 spectrum (Figure 7-4) suggests relative dissociation rates into the diketopiperazine and oxazolone pathways of 2:1 for dehydration of protonated AsnAla. Although these dissociation pathways are analogous to those of protonated AlaAsn, the relative rates into the two bifurcating pathways are significantly different.

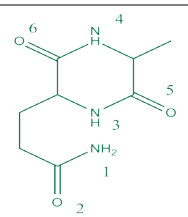
7.3.3. Dehydrated Alanine-Glutamine $[AlaGln+H - H_2O]^+$

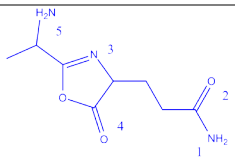
Six possible reaction pathways for the dehydration of protonated AlaGln are indicated in Scheme 7- 1, leading to five potential structures for the dehydrated ion shown in Tables 7-3 and Table S3 in ref. [129]. Nucleophilic attack of the side chain oxygen on the C-terminus and nucleophilic attack of the C-terminal oxygen on the side chain both lead to **13**.

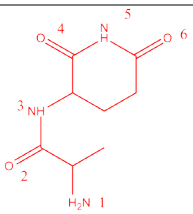
IRMPD of $[AlaGln+H–H_2O]^+$ induces only one fragment at 17 mass units below the precursor ion. Although analogy with the Asn-containing systems above would suggest ammonia-loss to be evidence for a diketopiperazine structure, the IR spectrum recorded is typical for an oxazolone fragment. Figure 7-5 shows the experimental IRMPD spectrum (black) compared with calculated IR spectra for the global minimum structure **10.6** (green, top left) and for the oxazolone alternatives **11.3** (dark blue, bottom left) and **11.5** (light grey, bottom left). The band near 1950 cm^{-1} is typical for an oxazolone moiety. A splitting of this band is clearly observed, suggesting the co-existence of two tautomers: **11.3** with protonation on the oxazolone nitrogen and **11.5** with protonation at the N-terminus, the latter being only 6 kJ/mol more stable than the former [49,120]. For **11.3**, the band at 1650 cm^{-1} corresponds to N-H bending at the oxazolone ring and that at 1665 cm^{-1} to N-H bending at the N-terminus. For **11.5**, the bands at 1630 cm^{-1} and 1520 cm^{-1} are assigned to N-H bending vibrations at the N-terminus.

The good match of the experimental spectrum with the oxazolone (**11.3** and **11.5**) spectrum suggests that the diketopiperazine structure (**10.6**) is at best a minor contributor to the population, in contrast with the observations for the Asn

analogue. The band at 1755 cm^{-1} in the computed spectrum of **10.6** and the weak band near 1780 cm^{-1} in the observed spectrum that remains unexplained by the oxazolone structures may hint at such a small fraction of diketopiperazine, although both diketopiperazine and oxazolone structures would have to decay by NH_3-loss in this case.







Name → ↓ $[H]^+$ Site	10 Diketopiperazine	11 Oxazolone	12 Glutarimide
1	+ 92	+ 123	+ 13
2	[a]	+ 132	+ 27
3	+ 67	+ 48	+ 119
4	+ 110	+ 155	[b]
5	+ 35	+ 42	+ 201
6	**0**		+ 146

Table 7-3. *Calculated relative Gibbs free energies in kJ/mol for potential ion structures resulting from H_2O loss from protonated AlaGln, each given for different potential protonation sites.*

[a] Proton migrates to position 6 during optimization.

[b] Proton migrates to position 2 during optimization.

The right panel of Figure 7-5 shows the IRMPD spectrum of the MS^3 fragment ion $[AlaGln+H–H_2O–NH_3]^+$ at m/z 183. Comparison with theoretical spectra suggests that NH_3 is detached from the oxazolone structure at the N-terminus. The band at 1835 cm^{-1} is due to C=O stretching in the oxazolone ring, that at 1685 cm^{-1} is attributed to C=O stretching in the side chain and the band at 1600 cm^{-1} is due to N-H bending in the side chain.

Scheme 7-2c summarizes the reaction mechanism. As for the Asn-containing dipeptides, water is lost from the C-terminus, but in this case it leads exclusively to an oxazolone structure. Unlike the oxazolone structures formed from the Asn-containing dipeptides, secondary fragmentation of the oxazolone ion involves the elimination of NH_3 rather than CO.

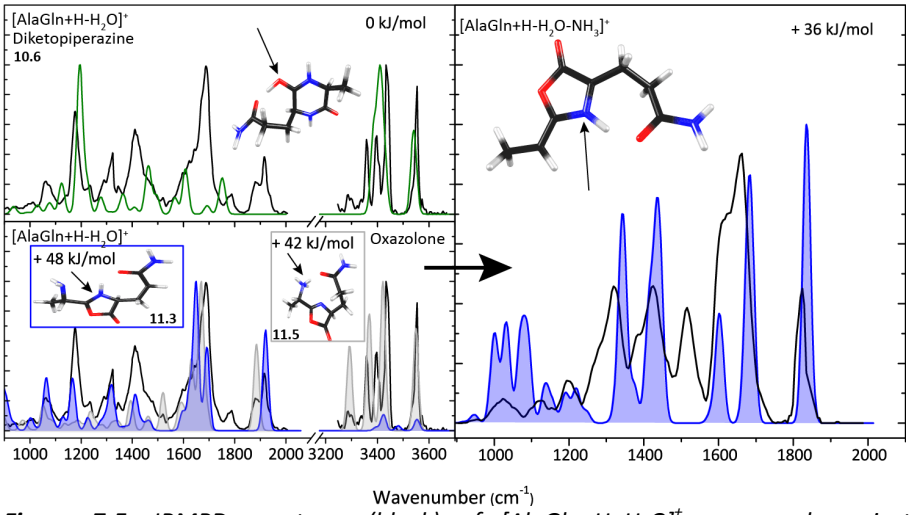
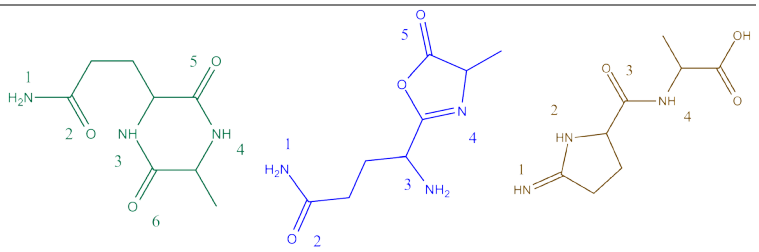


Figure 7-5. *IRMPD spectrum (black) of $[AlaGln+H–H_2O]^+$ compared against computed spectra for diketopiperazine **10.6** (top left) and oxazolones **11.3** and **11.5** (bottom left). The IRMPD spectrum for the $[AlaGln+H–H_2O–NH_3]^+$ MS^3 fragment ion (right) provides further evidence for an oxazolone structure.*

7.3.4. Dehydrated Glutamine-Alanine, [GlnAla + H – H_2O]$^+$

Peptides having Glu or Gln as their N-terminal residue are known for their deviating reaction mechanisms [3,8,9,63,137]. In particular, NH_3-loss from peptides with an N-terminal Gln residue and H_2O loss from peptides with an N-terminal Glu residue has been shown to form fragments with a pyroglutamyl moiety at the N-terminus [3,9,137]. Nucleophilic attack by the N-terminal nitrogen onto the side-chain δ-carbon atom is in both cases the *modus operandi*. Moreover, it has been previously suggested that doubly-protonated peptide ions that contain glutamine at the N-terminus and a basic or neutral amino acid at the C-terminus eliminate an H_2O molecule from the Gln side chain [3,9,137].



Name →	15	16	17
↓ $[H]^+$ Site	Diketopiperazine	Oxazolone	Imino-proline
1	+ 86	+ 206	0
2	[a]	[b]	+ 99
3	+ 152	+ 54	+ 146
4	+ 118	+ 91	+ 186
5	+ 83	+ 241	
6	+ 35		

Table 7-4. *Calculated relative Gibbs free energies in kJ/mol for possible structures resulting from H_2O loss from protonated GlnAla, each given for different protonation sites.*

[a] Proton migrates to position 6 during optimization.

[b] Proton migrates to position 3 during optimization.

However, the same singly charged peptides were reported to eliminate NH_3 [3,9]. The mechanism proposed for the reaction is a nucleophilic attack of the N-terminal nitrogen on the glutamine side chain leading to the loss of an H_2O molecule [3]. This mechanism follows the brown arrow in Scheme 7-1 and leads to the formation of structure **17**, which we shall designate by its imino-substituted proline residue. Here we follow the loss of H_2O from singly charged GlnAla.

Table 7-4 as well as Table S4 in in ref. [129]. list the five structures resulting from the reaction mechanisms proposed on the bottom row of Scheme 7-1. Ion **17.1**, the imino-prolinyl structure protonated on the imino nitrogen, is the global minimum. Structures **15** and **16** are the isomeric b_2-type ions resulting from H_2O loss from the C-terminus. Ions **18** and **19** are alternative structures resulting from a nucleophilic attack on the side chain and are not b-type ions.

Figure 7-6 compares the experimental spectrum of $[GlnAla+H–H_2O]^+$ with the computed spectrum for imino-prolynyl structure **17.1**. From the favorable agreement between experimental and theoretical spectra particularly in the 1000 – 2000 cm^{-1} range, we conclude that in contrast to the other dipeptides in this study, protonated GlnAla eliminates H_2O from the Gln side chain and not from the C-terminus. Further support for this assignment comes from the computed spectra for diketopiperazine (**15**) and oxazolone (**16**) structures, which show poor agreement with experiment, see Figure S4 in ref. [129]. Moreover, the presence of a carboxylic acid group in **17**, but not in **15** and **16**, can explain the observation of H_2O loss as the MS^3 channel, which was not observed for the other dipeptides in this study. As Table 7-4 indicates, the Gibbs free energies of the oxazolone and diketopiperazine structures are at least 54 and 35 kJ/mol higher than the imino-prolinyl structure, respectively.

In the computed spectrum of **17.1**, the band around 1730 cm^{-1} corresponds to unresolved carboxyl and amide C=O stretches and the band at 1684 cm^{-1} is due to NH_2 bending. The band at 1506 cm^{-1} is due to NH bending of the peptide bond. The spectrum in hydrogen stretching range matches poorly with the computation, but we suspect that this is due to threshold effects, where excitation with the OPO laser is unable to reach up to the dissociation threshold at most frequencies.

Further support for our structural assignment is provided by analysis of the subsequent MS^3 fragments. Secondary CID channels are loss of H_2O and H_2O+CO. The second H_2O is eliminated from the C-terminus leading to the formation of an oxazolone moiety, giving the structure shown in the lower left panel of Figure 7-6. The computed IR spectrum for this structure is seen to match fairly well with the experimental spectrum for this MS^3 ion in terms of frequencies, although relative intensities deviate somewhat. The band at 1880 cm^{-1} is due to the oxazolone C=O stretch and the peak at 1697 cm^{-1} is due to the CN stretch of the protonated imine moiety. The predicted spectrum for the $[GlnAla+H–H_2O–CO]^+$ MS^3 ion is also in fair agreement with the experiment in terms of frequencies, although again relative intensities deviate substantially. We suspect that threshold effects in the IRMPD process are at the origin of these deviations. The band at 1800 cm^{-1} is due to the C=O stretch and that at 1690 cm^{-1} to a bending vibration of the amino group.

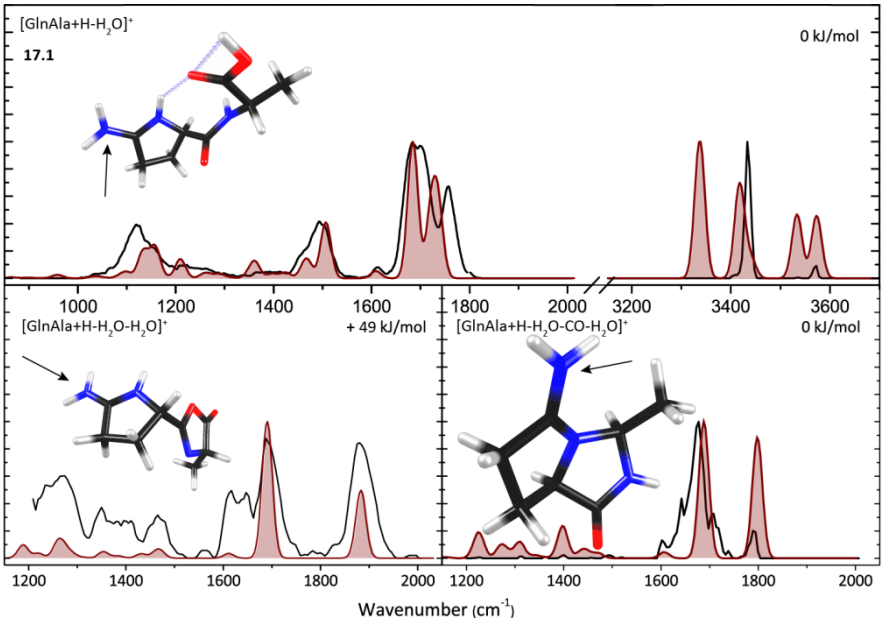


Figure 7-6. *Experimental IRMPD spectrum of the $[GlnAla+H-H_2O]^+$ MS^2 ion (top) and the $[GlnAla+H-H_2O-H_2O]^+$ MS^3 ion (bottom left) and $[GlnAla+H-H_2O-CO-H_2O]^+$ (bottom right) MS^3 ions, compared against the best matching computed spectra. The relatively poor match between experimental and computed spectra in the H-stretching range for the MS^2 fragment is tentatively attributed to the lower pulse energy of the OPO and/or a relatively high threshold to dissociation for this specific species, which would enhance non-linear effects and prevent excitation to above the threshold on some of the predicted resonances.*

Scheme 7-2d presents the suggested reaction mechanism for the dehydration of protonated GlnAla and the suggested pathway for sequential loss of H_2O and $CO+H_2O$. We conclude that nucleophilic attack occurs from the N-terminus on the glutamine side chain. Dehydration follows the same mechanism as observed for deamidation of the peptide [3,38,42,137]. Moreover, this mechanism is also analogous to deamidation of the Gln amino acid and the dehydration of the Glu amino acid (see chapter 3). After dehydration of protonated GlnAla, sequential fragmentation proceeds via two parallel mechanisms, one leading to loss of H_2O from the C-terminus and the other leading to loss of H_2O and CO.

7.4. Conclusion

Experimental infrared ion spectroscopy combined with quantum-chemical calculations has enabled us to assign the product ion structures resulting from H_2O loss from a series of protonated dipeptides containing Asn or Gln. As well, the structures of several MS^3 fragments were elucidated. From the assigned product ion structures we suggest the reaction mechanisms involved. Combined with our previous study on NH_3 loss from these peptides, [63] an extensive molecular structure map of the reaction channels after MS^n can be constructed as shown in Figure 7-7.

Our earlier study demonstrated that the deamidation reactions of these dipeptides take place from the Asn or Gln side chain. If we compare the reaction mechanisms assigned for the deamidation reaction with the mechanisms proposed here for the dehydration reaction, we see that only for protonated GlnAla is the mechanism analogous, where the same nucleophilic attack on the side chain takes place during both fragmentation reactions. This is in line with the dehydration reaction of aspartic acid and glutamic acid, where the side chain is attacked by the N-terminus (see chapter 3). The assigned structure for protonated GlnAla is not a b_2 ion structure but is in agreement with reported structures [3,9,137]. A diketopiperazine structure has also been found for the b_2 ion of AsnAlaAla [11,120]. For AlaAsnAla the formation of a succinimide structure (which involves the side chain) has been reported [11], but this structure was not identified in the dipeptides that contain asparagine in the work presented here. The assigned oxazolone structure for H_2O-loss from protonated AsnAla is in agreement with the oxazolone structure for AsnGly-H_2O previously proposed [59]. However, no bifurcating mechanism to both an oxazolone and diketopiperazine stucture after water loss was proposed [59]. Dehydration of protonated AlaAsn forms the b_2-ion and does not involve the Asn side chain, in contrast to b_2 formation from protonated AlaAsnAla, where the Asn side chain is involved in the formation of a succinimide structure [11]. This suggests that the identity of the third residue is of influence in the formation of b_2-type ions. Detailed transition state calculations are required to explain these observations, where the present structural identification may be used to guide such calculations.

On the basis of the calculated relative Gibbs free energies, the diketopiperazine structures are consistently the lowest energy structures. The observed oxazolone structures have higher relative free energies (48 - 73 kJ/mol), but are often suggested to form more likely because of kinetic arguments [65,120,121,124,138-140]. Here, we find that for the Asn-containing peptides, H_2O-loss produces both diketopiperazine and oxazolone structures.

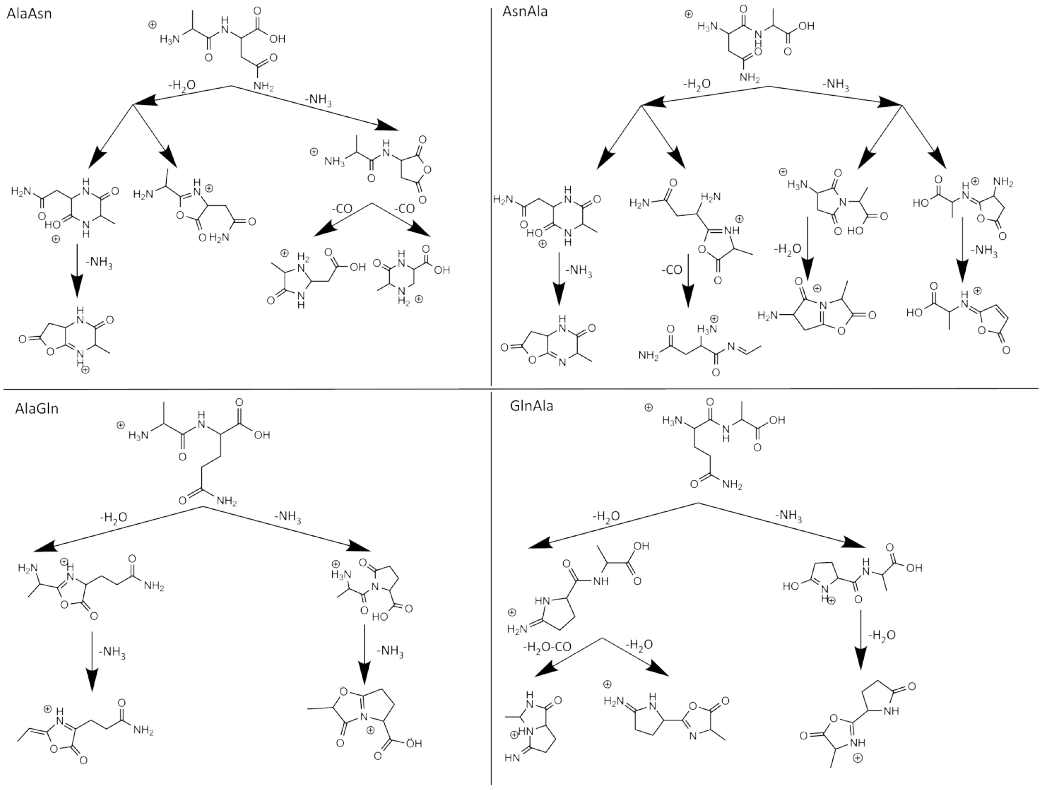


Figure 7-7. *Molecular structure map of deamidation and dehydration reactions in the Asn and Gln containing protonated dipeptides based on the spectroscopically identified fragment ion structures established.*

Bibliography

[1] Heaton, A.L., Armentrout, P.B., *J. Am. Chem. Soc.*, **130**, (2008), 10227.
[2] Orlowski, M., Meister, A., Enzymology of Pyrrolidone Carboxylic Acid, Academic Press, New York, 1971.
[3] Neta, P., Pu, Q., Kilpatrick, L., Yang, X., Stein, S.E., *J. Am. Soc. Mass. Spectrom.*, **18**, (2007), 27.
[4] Dunkelberger, E.B., Buchanan, L.E., Marek, P., Cao, P., Raleigh, D.P., Zanni, M.T., *J. Am. Chem. Soc.*, **134**, (2012), 12658.
[5] Robinson, N.E., Robinson, A.B., *Proc. Natl. Acad. Sci. U.S.A.*, **98**, (2001), 944.
[6] Harrison, A.G., Yalcin, T., *Int. J. Mass Spectrom. And Ion Processes*, **165-166**, (1997), 339.
[7] Heaton, A.L., Armentrout, P.B., *J. Am. Soc. Mass Spectrom.*, **20**, (2009), 852.
[8] Dookeran, N.N., Yalcin, T., Harrison, A.G., *J. Mass Spectrom.*, **31**, (1996), 500.
[9] Harrison, A.G., *J. Mass Spectrom.*, **38**, (2003), 174.
[10] Rogalewicz, F., Hoppilliard, Y., Ohanessian, G., *Int. J. Mass Spectrom. And Ion Processes*, **195/196**, (2000).
[11] Grzetic, J., Oomens, J., *J. Am. Soc. Mass Spectrom.*, **24**, (2013), 1228.
[12] Grzetic, J., Oomens, J., *Int. J. Mass Spectrom.*, **354-355**, (2013), 70.
[13] Paizs, B., Suhai, S., *Mass Spec. Rev.*, **24**, (2005), 508.
[14] Osburn, S., Burgie, T., Berden, G., Oomens, J., O'Hair, R.A.J., Ryzhov, V., *J. Phys. Chem. A.*, **177**, (2013), 1144.
[15] Martens, J.K., Grzetic, J., Berden, G., Oomens, J., *Int. J. of Mass Spectrom.*, **377**, (2015), 179-187.
[16] Heaton, A.L., Bowman, V.N., Oomens, J., Steill, J.D., Armentrout, P.B., *J. Phys. Chem.* , **113**, (2009), 5519.
[17] Campbell, S., Beauchamp, J.L., Rempe, M., Lichtenberger, D.L., *Int. J. Mass Spectrom. and Ion Processes*, **117**, (1992), 83.
[18] Dongre, A.R., Jones, J.L., Somogyi, A., Wysocki, V.H., *J. Am. Chem. Soc.*, **118**, (1996), 8365.
[19] Wysocki, V.H., Tsaprailis, G., Smith, L.L., Breci, L.A., *J. Mass Spectrom.*, **35**, (2000), 1399.
[20] Boyd, R., Somogyi, A., *J. Am. Soc. Mass Spectrom.*, **21**, (2010), 1275-1278.
[21] Konuklar, F.A., Aviyente, V., Sen, T.Z., Bahar, I., *J. Mol. Model*, **7**, (2001), 147.
[22] Heaton, A.L., Ye, S.J., Armentrout, P.B., *J. Phys. Chem. A.*, **112**, (2008), 3328.
[23] Oomens, J., Steill, J.D., Redlich, B., *J. Am. Chem. Soc.*, **131**, (2009), 4310.
[24] Patrick, A.L., Stedwell, C.N., Schindler, B., Compagon, I., Berden, G., Oomens, J., Polfer, N.C., *Int. J. Mass Spectrom.*, **379**, (2015), 26-32.
[25] Correia, C.F., Balaj, P.O., Scuderi, D., Maître, P., Ohanessian, G., *J. Am. Chem. Soc.* , **130**, (2008), 3359-3370.
[26] Oomens, J., Sartakov, B.G., Meijer, G., von Helden, G., *Int. J. Mass Spectrom.*, **254**, (2006), 1-19.

[27] Oepts, D., van der Meer, A.F.G., van Amersfoort, P.W., *Infrared Phys. Technol.*, **36**, (1995), 297.
[28] Valle, J.J., Eyler, J.R., Oomens, J., Moore, D.T., van der Meer, A.F.G., von Helden, G., Meijer, G., Hendrickson, C.L., Marshall, A.G., Blakney, G.T., *Rev. Sci. Instrum.*, **76**, (2005), 023103.
[29] Marshall, A.G., Wang, T.C.L., Ricca, T.L., *J. Am. Chem. Soc.*, **107**, (1985), 7893.
[30] Polfer, N.C., Oomens, J., *Phys.Chem.Chem.Phys.*, **9**, (2007), 3804.
[31] Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G.A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H.P., Izmaylov, A.F., Bloino, J., Zheng, G., Sonnenberg, J.L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J., J.A., Peralta, J.E., Ogliaro, F., Bearpark, M., Heyd, J.J., Brothers, E., Kudin, K.N., Staroverov, V.N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Rega, N., Millam, N.J., Klene, M., Knox, J.E., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Martin, R.L., Morokuma, K., Zakrzewski, V.G., Voth, G.A., Salvador, P., Dannenberg, J.J., Dapprich, S., Daniels, A.D., Farkas, O., Foresman, J.B., Ortiz, J.V., Cioslowski, J., Fox, D.J., (2009).
[32] Case, D.A., Darden, T., Cheatham III, T.E., Simmerling, C., Wang, J., Duke, R.E., Luo, R., Walker, R.C., Zhang, W., Merz, K.M., Roberts, B.P., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Goetz, A.W., Kolossváry, I., Wong, K.F., Paesani, F., Vanicek, J., Wolf, R.M., Liu, J., Wu, X., Brozell, S.R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D.R., Mathews, D.H., Seetin, M.G., Salomon-Ferrer, R., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., Kollman, P.A., (2012).
[33] Harrison, A.G., Mercer, R.S., Reiner, E.J., Young, A.B., Boyd, R.K., March, R.E., Porter, C.J., *Int. J. Mass Spectrom. and Ion Processes.*, **74**, (1986), 13.
[34] Harrison, A.G., Tu, Y., *J. Mass Spectrom.*, **33**, (1998), 532-542.
[35] Drayß, M.K., Blunk, D., Oomens, J., Gao, B., Wyttenbach, T., Bowers, M.T., Schäfer, M., *J. Phys. Chem. A.*, **113**, (2009), 9543.
[36] Asmis, K.R., Pivonka, N.L., Santambrogio, G., Brümmer, M., Kaposta, C., Neumark, D.M., Wöste, L., *Science*, **299**, (2003), 1375.
[37] Harrison, A.G., *Int. J. Mass Spectrom.*, **210-211**, (2001), 361.
[38] Kempkes, L.J.M., Martens, J.K., Grzetic, J., Berden, G., Oomens, J., *Rapid Commun. Mass Spectrom.*, **30**, (2016), 483-490.
[39] Jonsson, A.P., Bergman, T., Jornvall, H., Griffiths, W.J., *J. Am. Chem. Soc*, **12**, (2001), 337-342.
[40] Geiger, T., Clarke, S., *J. Biol. Chem.*, **262**, (1987), 785-794.
[41] Lindner, H., Helliger, W., *Exp. Geront.*, **36**, (2001), 1551-1563.
[42] Baldwin, M.A., Falick, A.M., Gibson, B.W., Prusiner, S.B., Stahl, N., Burlingame, A.L., *J. Am. Soc. Mass Spectrom.*, **1**, (1990), 258-264.

[43] Han, X., Aslanian, A., Yates, J.R., *Curr. Opin. Chem. Biol.*, **12**, (2008), 483-490.
[44] Lucas, B., Grégoire, G., Lemaire, J., Maître, P., Ortega, J., Rupenyan, A., Reimann, B., Schermann, J.P., Desfrançois, C., *Phys. Chem. Chem. Phys.*, **6**, (2004), 2659-2663.
[45] Oomens, J., Young, S., Molesworth, S., van Stipdonk, M., *J. Am. Soc. Mass Spectrom.*, **20**, (2009), 334-339.
[46] Fung, Y.M.E., Besson, T., Lemaire, J., Maître, P., Zubarev, R.A., *Angew. Chem.*, **121**, (2009), 8490-8492.
[47] Perkins, B.R., Chamot-Rooke, J., Yoon, S.H., Gucinski, A.C., Somogyi, A., Wysocki, V.H., *J. Am. Chem. Soc*, **131**, (2009), 17528-17529.
[48] Sinha, R.K., Erlekam, U., Bythell, B.J., Paizs, B., Maitre, P., *J. Am. Soc. Mass Spectrom.*, **22**, (2011), 1645-1650.
[49] Polfer, N.C., Oomens, J., Suhai, S., Paizs, B., *J. Am. Soc.*, **129**, (2007), 5887-5897.
[50] Polfer, N.C., Oomens, J., *Mass Spec. Rev.*, **28**, (2009), 468-494.
[51] Bythell, B.J., Maître, P., Paizs, B., *J. Am. Chem. Soc.*, **132**, (2010), 14766-14779.
[52] Bythell, B.J., Erlekam, U., Paizs, B., Maître, P., *ChemPhysChem*, **10**, (2009), 883-885.
[53] Erlekam, U., Bythell, B.J., Scuderi, D., van Stipdonk, M., Paizs, B., P.;, M., *J. Am. Chem. Soc*, **131**, (2009), 11503-11508.
[54] Farrugia, J.M., O'Hair, R.A.H., Reid, G., *Int. J. Mass Spectrom.*, **210/211**, (2001), 71-87.
[55] Lee, Y.J., Lee, Y.M., *Rapid Commun. Mass Spectrom.*, **18**, (2004), 2069-2076.
[56] Winter, D., Seidler, J., Hahn, B., Lehmann, W.D., *J. Am. Soc. Mass Spectrom.*, **21**, (2010), 1814-1820.
[57] Wu, R., McMahon, T.B., *Chem.Phys.Chem.*, **9**, (2008), 2826.
[58] Bleiholder, C., Paizs, B., *Theor. Chem. Acc.*, **125**, (2010), 387-396.
[59] Boles, G.C., Wu, R.R., Rodgers, M.T., Armentrout, P.B., *J. Phys. Chem. B*, **120**, (2016), 6525-6545.
[60] Martens, J., Grzetic, J., Berden, G., Oomens, J., *Nature Commun.*, **7**, (2016), 11754.
[61] Amster, J., *J. Mass Spectrom.*, **31**, (1996), 1325-1337.
[62] Marshall, A.G., Hendrickson, C.L., Jackson, G.S., *Mass Spec. Rev.*, **17**, (1998), 1-35.
[63] Kempkes, L.J.M., Martens, J., Grzetic, J., Berden, G., Oomens, J., *J. Am. Soc. Mass Spectrom.*, **27**, (2016), 1855-1869.
[64] Barlow, C.K., O'Hair, R.A.H., *J. Mass Spectrom.*, **43**, (2008), 1301-1319.
[65] Armentrout, P.B., Heaton, A.L., *J. Am. Soc. Mass Spectrom.*, **23**, (2012), 621-631.
[66] Ortiz, D., Martin-Gago, P., Riera, A., Song, K., Salpin, J.Y., Spezia, R., *Int. J. Mass Spectrom.*, **335**, (2013), 33-44.
[67] Bischoff, R., Kolbe, H.V.J., *J. Chromatogr. B*, **662**, (1994), 261-278.
[68] Guan, X., Wang, B., Wang, H., Liu, J., Li, Y., Guo, X., *Rapid Commun. Mass Spectrom.*, **31**, (2017), 649-657.

[69] Nilsson, M.R., Driscoll, M., Raleigh, D.P., *Protein Sci.*, **11**, (2002), 342-349.
[70] Robinson, N.E., Robinson, A.B., *Proc. Natl. Acad. Sci. U.S.A.*, **98**, (2001), 12409-12413.
[71] Boles, G.C., Owen, C.J., Berden, G., Oomens, J., Armentrout, P.B., *Phys. Chem. Chem. Phys.*, **19**, (2017), 12394-12406.
[72] Ye, S.J., Armentrout, P.B., *J. Phys. Chem. A*, **112**, (2008), 3587-3596.
[73] Mookherjee, A., Van Stipdonk, M.J., Armentrout, P.B., *J. Am. Soc. Mass Spectrom.*, **28**, (2017), 739-757.
[74] Ervin, K.M., Armentrout, P.B., *J. Chem. Phys.*, **83**, (1985), 166-189.
[75] Muntean, F., Armentrout, P.B., *J. Chem. Phys.*, **115**, (2001), 1213-1228.
[76] Armentrout, P.B., *J. Am. Soc. Mass Spectrom.*, **13**, (2002), 419-434.
[77] Moision, R.M., Armentrout, P.B., *J. Am. Soc. Mass Spectrom.*, **18**, (2007), 1124-1134.
[78] Kim, T., Tolmachev, A.V., Harkewicz, R., Prior, D.C., Anderson, G., Udseth, H.R., Smith, R.D., *Anal. Chem*, **72**, (2000), 2247-2255.
[79] Carl, D.R., Chatterjee, B.K., Armentrout, P.B., *J. Chem. Phys.*, **132**, (2010), 44303.
[80] Carl, D.R., Moision, R.M., Armentrout, P.B., *J. Am. Soc. Mass Spectrom.*, **20**, (2009), 2312-2317.
[81] Chen, Y., Rodgers, M.T., *J. Am. Chem. Soc*, **134**, (2012).
[82] Carpenter, J.E., McNary, C.P., Furin, A., Sweeney, A.F., Armentrout, P.B., *J. Am. Soc. Mass Spectrom.*, **28**, (2017), 1876-1888.
[83] Teloy, E., Gerlich, D., *Chem. Phys.*, **4**, (1974), 417-427.
[84] Gerlich, D., *Adv. Chem. Phys.*, **82**, (1992), 1-176.
[85] Aristov, N., Armentrout, P.B., *J. Phys. Chem.*, **90**, (1986), 5135-5140.
[86] Dalleska, N.F., Honma, K., Sunderlin, L.S., Armentrout, P.B., *J. Am. Chem. Soc.*, **116**, (1994), 3519-3528.
[87] Daly, N.R., *Rev. Sci. Instrum.*, **31**, (1960), 264-267.
[88] Beyer, T.S., Swinehart, D.F., *Commun. ACM*, **16**, (1973), 379.
[89] Stein, S.E., Rabinovich, B.S., *Chem. Phys. Lett.*, **49**, (1977), 183-188.
[90] Stein, S.E., Rabinovitch, B.S., *J. Chem. Phys.*, **58**, (1973), 2438-2445.
[91] Gilbert, R.G., Smith, S.C., Theory of Unimolecular and Recombination Reactions, Blackwell Scientific Publications, 1990.
[92] Robinson, P.J., Holbrook, K.A., Unimolecular Reactions, Wiley Interscience, New York, 1972.
[93] Rodgers, M.T., Armentrout, P.B., *J. Chem. Phys.*, **109**, (1998), 1787-1800.
[94] Rodgers, M.T., Ervin, K.M., Armentrout, P.B., *J. Chem. Phys.*, **106**, (1997), 4499-4508.
[95] Foresman, J.B., Frisch, A.E., Exploring Chemistry with Electronic Structure Methods, 2nd ed. ed., Gaussian, Inc., Pittsburgh, PA, 1996.
[96] Kempkes, L.J.M., Boles, G.C., Martens, J., Berden, G., Armentrout, P.B., Oomens, J., *J. Phys. Chem. A.*, **122**, (2018), 2424-2436.

[97] Boles, G.C., Wu, R.R., Rodgers, M.T., Armentrout, P.B., *J. Am. Soc. Mass Spectrom.*, **29**, (2018), 2341-2359.
[98] Martin-Somer, A., Martens, J., Grzetic, J., Hase, W.L., Oomens, J., Spezia, R., *J. Phys. Chem. A.*, **122**, (2018), 2612-2625.
[99] Spezia, R., Lee, S.B., Cho, A., Song, K., *Int. J. Mass Spectrom.*, **392**, (2015), 125-138.
[100] Meroueh, S.O., Wang, Y., Hase, W.L., *J. Phys. Chem. A.*, **106**, (2002), 9983-9992.
[101] Homayoon, Z., Pratihar, S., Dratz, E., Snider, R., Spezia, R., Barnes, G.L., Macaluso, V., Martin Somer, A., Hase, W.L., *J. Phys. Chem. A.*, **120**, (2016), 8211-8227.
[102] Spezia, R., Martens, J., Oomens, J., Song, K., *Int. J. Mass Spectrom.*, **388**, (2015), 40-52.
[103] Aebersold, R., Mann, M., *Nature*, **422**, (2003), 198-207.
[104] Steen, H., Mann, M., *Nat. Rev. Mol. Cell Biol.*, **5**, (2004), 699-711.
[105] Hase, W.L., Buckowski, D.G., *Chem. Phys. Lett.*, **74**, (1980), 284-287.
[106] Verlet, C., *Phys. Rev.*, **159**, (1967), 98-103.
[107] Rocha, G.B., Freire, R.O.l., Simas, A.M., Stewart, J.J.P., *J. Comput. Chem.*, **27**, (2006), 1101-1111.
[108] Barnes, G.L., Hase, W.L., *J. Am. Chem. Soc.*, **131**, (2009), 17185-17193.
[109] Barnes, G.L., Young, K., Yang, L., Hase, W.L., *J. Chem. Phys.*, **134**, (2011), 0941061-09410613.
[110] Homayoon, Z., Macaluso, V., Martin-Somer, A., Nicola-Barbosa-Muniz, M.C., Borges, I., Hase, W.L., Spezia, R., *Phys. Chem. Chem. Phys.*, **20**, (2018), 3614-3629.
[111] Macaluso, V., Homayoon, Z., Spezia, R., Hase, W.L., *Phys. Chem. Chem. Phys.*, **20**, (2018), 19744-19749.
[112] Hase, W.L., Duchovic, R.J., Hu, X., Komornicki, A., Lim, K.F., Lu, D.H., Peslherbe, G.H., Swamy, K.N., Vande-Linde, S.R., Zhu, L., Varandas, A., Wang, H., Wolf, R.J., *QCPE Bull.*, **16**, (1996), 671.
[113] Stewart, J.J.P., Fiedler, L.J., Zhang, P., Zheng, J., Rossi, I., Hu, W.-P., Lynch, G.C., Liu, Y.-P., Chuang, Y.-Y., Pu, J., Li, J., Cramer, C.J., Fast, P.L., Truhlar, D.G., *Department of Chemistry and Supercomputing Institute, University of Minnesota: Minneapolis, MN*, (2015).
[114] Jeanvoine, Y., Largo, A., Hase, W.L., Spezia, R., *J. Phys. Chem. A.*, **122**, (2018), 869-877.
[115] Spezia, R., Martin-Somer, A., Macaluso, V., Homayoon, Z., Pratihar, S., Hase, W.L., *Faraday Discuss.*, **195**, (2016), 599-618.
[116] Li, W., Song, C., Bailey, D.J., Tseng, G.C., Coon, J.J., Wysocki, V.H., *Anal. Chem.*, **83**, (2011), 9540-9545.
[117] Aebersold, R., Goodlett, D.R., *Chem. Rev.*, **101**, (2001), 269-296.
[118] Yates, J.R., *J. Mass Spectrom.*, **33**, (1998), 1-19.
[119] Roepstorff, P., Fohlman, J., *Biom. Mass Spectrom.*, **11**, (1984), 601.

[120] Morrison, L.J., Chamot-Rooke, J., Wysocki, V.H., *Analyst.*, **139**, (2014), 2137-2143.
[121] Polfer, N.C., Oomens, J., Suhai, S., Paizs, B., *J. Am. Chem. Soc.*, **127**, (2005), 17154-17155.
[122] Frison, G., van der Rest, G., Turecek, F., Besson, T., Lemaire, J., Maître, P., ;, Chamot-Rooke, J., *J. Am. Chem. Soc.*, **130**, (2008), 14916-14917.
[123] Yoon, S.H., Chamot-Rooke, J., Perkins, B.R., Hilderbrand, A.E., Poutsma, J.C., Wysocki, V.H., *J. Am. Chem. Soc.*, **130**, (2008), 17644-17645.
[124] Balta, B., Aviyente, V., Lifshitz, C., *J. Am. Chem. Soc.*, **14**, (2003), 1192-1203.
[125] Ballard, K.D., Gaskell, S.J., *J. Am. Soc. Mass Spectrom.*, **4**, (1993), 477-481.
[126] Reid, G.E., Simpson, R.J., O'Hair, R.A.J., *J. Am. Soc. Mass Spectrom.*, **9**, (1998), 945-956.
[127] Zou, S., Oomens, J., Polfer, N.C., *Int. J. of Mass Spectrom.*, **316-318**, (2012), 12-17.
[128] Verkerk, U.H., Zhao, M.J., van Stipdonk, M.J., Bythell, B.J., Oomens, J., Hopkinson, A.C., Siu, K.W.M., *J. Phys. Chem. A.*, **115**, (2011), 6683-6687.
[129] Kempkes, L.J.M., Martens, J., Berden, G., Oomens, J., *Int. J. Mass Spectrom.*, **429**, (2018), 90-100.
[130] Headrick, J.M., Diken, E.G., Walters, R.S., Hammer, N.I., Christie, R.A., Cui, J., Myshakin, E.M., Duncan, M.A., Johnson, M.A., Jordan, K.D., *Science*, **308**, (2005), 1765-1769.
[131] Hurtado, P., Gamez, F., Hamad, S., Martinez-Haya, B., Steill, J.D., Oomens, J., *J. Phys. Chem. A.*, **115**, (2011), 7275-7282.
[132] Li, X., Oomens, J., Eyler, J.R., Moore, D.T., Iyengar, S.S., *J. Chem. Phys.*, **132**, (2010), 244301.
[133] Headrick, J.M., Bopp, J.C., Johnson, M.A., *J. Chem. Phys.*, **121**, (2004), 11523-11526.
[134] Fridgen, T.D., MacAleese, L., Maitre, P., McMahon, T.B., Boissel, P., Lemaire, J., *Phys. Chem. Chem. Phys.*, **7**, (2005), 2747-2755.
[135] Fridgen, T.D., McMahon, T.B., MacAleese, L., Lemaire, J., Maitre, P., *J. Phys. Chem. A.*, **108**, (2004), 9008-9010.
[136] Wysocki, V.H., Resing, K.A., Zhang, Q., Cheng, G., *Methods*, **35**, (2005), 211-222.
[137] Godugu, B., Neta, P., Simon-Manso, Y., Stein, S.E., *J. Am. Soc. Mass Spectrom.*, **21**, (2010), 1169-1176.
[138] Yalcin, T., Csizmadia, I.G., Peterson, M.R., Harrison, A.G., *J. Am. Soc. Mass Spectrom.*, **7**, (1996), 233-242.
[139] Nold, M.J., Wesdemiotis, C., Yalcin, T., Harrison, A.G., *Int. J. Mass Spectrom. and Ion Processes*, **164**, (1997), 137-153.
[140] Paizs, B., Suhai, S., *Rapid Commun. Mass Spectrom.*, **16**, (2002), 375-389.

Part III – Spectroscopic investigation of electron transfer dissociation product ions

Chapter 8

Spectroscopic characterization of an extensive set of c-type peptide fragment ions formed by electron transfer dissociation suggests exclusive formation of amide isomers

---

Electron attachment dissociation (ECD, ETD) applied to gaseous multiply protonated peptides leads predominantly to backbone N – C_α bond cleavages and the formation of *c*- and *z*-type fragment ions. The mechanisms involved in the formation of these ions have been the subject of much discussion. Here, we determine the molecular structures of an extensive set of *c*-type ions produced by ETD using infrared ion spectroscopy. Nine c_3- and c_4-ions are investigated to establish their C-terminal structure as either enol-imine or amide isomers by comparison of the experimental IR spectra with quantum-chemically predicted spectra for both structural variants. The spectra suggest that all *c*-ions investigated possess an amide structure; the absence of an NH bending mode at approximately 1000-1200 cm^{-1} serves as an important diagnostic feature.

* Adapted from: Kempkes, L.J.M.; Martens, J.; Berden, G.; Oomens, J., Spectroscopic characterization of c-type ions formed by Electron Transfer Dissociation of a Comprehensive set of tetra- and heptapeptides shows exclusive amide formation. *J. Phys. Lett.* **2018**, *9*, 6404-6411.

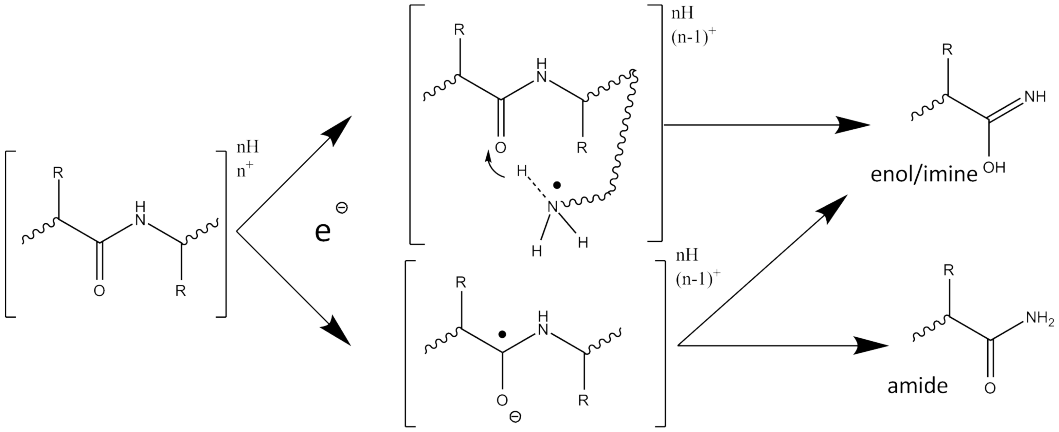
8.1. Introduction

Mass spectrometry (MS) is well established as the leading technique for protein sequencing.[1] Collision Induced Dissociation (CID) has long been the main tandem MS (MS/MS) method used to induce peptide fragmentation, cleaving protonated peptides at their amide bonds[2] and generating predominantly *b*- and *y*-type sequence ions.[3] However, the inability of CID MS/MS to sequence labile post translational modifications (PTMs) and its limited applicability in top-down protein sequencing[4-9] have led to an increase in the use of fragmentation strategies based on electron attachment to the multiply protonated target, in particular electron capture dissociation (ECD) and electron transfer dissociation (ETD).[10] ECD and ETD, jointly referred to as ExD, increase sequence coverage and thus enable top-down sequencing, mitigating the need for enzymatic digestion of a protein into smaller peptides[5,11-24], and moreover leave labile PTMs attached, which was originally suggested to be due to a non-ergodic nature of the dissociation process, but this hypothesis was later rejected.[25-27] ExD of multiply-charged gaseous proteins results mainly in *c*- and *z*-type ions, cleaving the backbone at the N – C_α bond.[24,28-35] Complementary information is obtained by combining CID and ExD for protein identification as different ion types are formed.

In ETD, an anionic species (often the radical anion of fluoranthene[10]) is guided into the ion trap and stored along with the multiply protonated precursor peptide ion of interest.[11,24,36] An ion/ion reaction results in electron transfer and charge-reduction of the peptide ion forming a radical species.[37] The charge recombination induces cleavage of the peptide backbone predominantly at the N – C_α bonds[24,28-35], resulting in *c*- and radical *z*-type sequence ions.

The precise reaction mechanisms of ETD have been the subject of extensive discussion[17,24,27,29,30,38-52]. Several mechanisms have been proposed with the Cornell[17,29,39,53] and Utah-Washington[40,44-47,54-56] mechanisms being the most prominent ones. In the Cornell mechanism, electron attachment occurs at a protonated site that is hydrogen-bonded to a nearby carbonyl.[17] Hydrogen atom transfer from the now neutralized protonation site to the carbonyl group then induces N – C_α bond cleavage and the formation of *c*-type ions with an enol-imine structure at their C-terminal end (Scheme 1). In the Utah-Washington hypothesis, the electron is captured in the π^*-orbital of an amide carbonyl H-bonded to a protonated site, producing a charge-stabilized amide anion-radical intermediate, which isomerizes by proton transfer to the peptide bond amide oxygen or

nitrogen;[34,57] proton transfer to the amide nitrogen upon cleavage of the N – C_α bond forms *c*-type ions with an amide moiety at the C-terminus. Charge stabilization by protonation, rather than by metal-ion coordination, was recently suggested to enhance the efficiency of *c*- and *z*-type ion formation.[52]



Scheme 8-1. *Schematic representation of two isomeric c-type product ions that have been proposed to be formed upon ETD. The top structure is the enol-imine c-type structure proposed to result from the Cornell mechanism. Following the Utah-Washington mechanism, c-type product ions can form as either an amide or enol-imine.*

Since the two mechanisms lead to two different (isomeric) product ions, with either an enol-imine or an amide terminus, identification of the molecular structure of the *c*-type ions may shed light on their mechanism of formation. The combination of mass spectrometry, ion spectroscopy and quantum-chemical calculations has become a powerful method for distinguishing isomers, tautomers and conformers of MS/MS reaction products,[58-63] which has previously been used for the structural characterization of an ECD-generated c_0-ion of a derivatized peptide.[64] Its structure was established to include an amide moiety as its C-terminus. More recently, the c_4-ion produced by ETD on the doubly protonated GL*GGK peptide was also assigned as an amide structure on the basis of UV photodissociation and IRMPD spectroscopy experiments along with density functional theory (DFT) calculations. Calculated spectra distinguished the amide tautomer from the enol-imine tautomer by strong IR bands of the enol-imine product ion at 1040-1050 cm^{-1} and 1190-1220 cm^{-1}, which were absent in the experimental spectrum.[65]

Ion mobility has also been used to characterize ETD fragment ions (c_3, c_4, z_3 and z_4 of $[AAHAL+2H]^{2+}$).[66] Although the collisional cross sections (Ω) showed close agreement with calculations, the values for enol-imine and amide tautomers were nearly identical not allowing the authors to distinguish between them. In another study, the CID fragmentation pattern of the c_5-ion of GAILKGAILR was found to be essentially identical to that of the synthesized $[GAILK\text{-}NH_2+H]^+$ analog.[67] In contrast, CID fragmentation of the intact peptide $[GAILK+H]^+$ showed a y_3-fragment ion and several neutral loss-species, which were absent in the CID spectrum of the c_5-ion (and of $[GAILK\text{-}NH_2+H]^+$).

Here we address the question of whether the formation of amide *c*-type ions is generic or not, by applying infrared multiple photon dissociation (IRMPD) spectroscopy to probe the structure of an extensive set of ETD *c*-type ions. This also addresses the question of whether the fragment ion structure is influenced by the identity of the amino acid residue at the cleavage site. Peptides containing a Lys residue at the N-terminus were selected to favor *c*-ion formation upon ETD. The peptide length is varied (tetra- and heptapeptides), as is the amino acid residue N-terminal to cleavage site (Ala, His or Phe); a peptide with a His residue C-terminal to the cleavage site was also included. An IRMPD spectrum was also recorded for the NH_2 capped peptide of AAAA, as reference for an amide-structure *c*-type ion.

8.2. Experimental and computational methods

8.2.1. IRMPD spectroscopy

The experiments made use of a modified ion trap tandem mass spectrometer (Bruker AmaZon ETD Speed) coupled to the beamline of the IR free electron laser FELIX[58]. Doubly protonated peptide ions were generated using electrospray ionization (ESI) from 10^{-6} M solutions in 50:50 acetonitrile:water with ~0.5% formic acid. The doubly charged precursor ions of interest were mass isolated in the quadrupole ion trap. Electron transfer dissociation (ETD) was effected by admitting fluoranthene radical anions[37] to the trap reacting with the stored peptide cations for 300 ms. The basic Lys residue in the first position generates a prominent series of *c*-type ions, out of which the singly charged c_3 or c_4 ion of interest was mass isolated. Note that Lys is preferred over other basic residues such as Arg or His because of its relative silence in the IR spectrum, in contrast with the guanidinium and imidazolium side chains which feature strong IR absorptions overlapping with, and therefore obscuring, the diagnostic amide I and II features in the spectra. A

potential side-effect of this choice is the alleged catalytic behavior of the ammonium group in the conversion of enol-imine to amide structures.[68]

The *c*-type fragment ion was then irradiated with two IR pulses of the FELIX free-electron laser.[58,65] FELIX produced 6 µs long macropulses of 20–60 mJ at a 10 Hz repetition rate having a bandwidth of ~0.5% of the center frequency. The IR induced dissociation yield, calculated as ΣI(fragment ions)/ΣI(parent + fragment ions), at each laser frequency was determined from five averaged mass spectra. Plotting the yield as a function of laser frequency then generates an infrared spectrum. The yield is linearly corrected for the frequency dependent pulse energy and the IR frequency is calibrated using a grating spectrometer.

8.2.2. Computational chemistry

For all *c*-ions, enol-imine and amide isomeric structures were optimized and their infrared spectra were calculated using density functional theory (DFT) at the B3LYP/6-31++ G(d,p) level using Gaussian 09 revision D01[69]. These spectra were used for qualitative initial comparison with the experimental spectra. A molecular mechanics/molecular dynamics (MM/MD) approach employing AMBER 12[70] was then used to further search for lower-energy conformers of both isomeric motifs. Within AMBER an initial MM geometry optimization was performed, followed by a simulated annealing procedure up to 273-500 K, resulting in 500 structures. These structures were grouped based on rms atom positions to give 20-30 candidate structures, which were optimized using DFT at the B3LYP/6-31++G(d,p) level. Their spectra were compared with the experimental spectra. Computed harmonic vibrational frequencies were scaled by 0.975 and convoluted with a 25 cm^{-1} full-width-at-half-maximum (FWHM) Gaussian line shape. Single-point electronic energies were calculated at the MP2/6-311+G(2d,2p) level using the B3LYP/6-31++G(d,p) optimized structures. The computational procedure is described in more detail elsewhere[58,59,71].

8.3. Results and discussion

Figure 8-1 presents the IRMPD spectrum of the c_3 ion of $[KAAA+2H]^{2+}$ compared with spectra predicted for different structural variants by DFT. The gray trace in the left panel is the computed spectrum for the lowest-energy conformer found for the amide isomeric form. The spectrum in blue represents an alternative, higher-energy (+18 kJ/mol) conformer of the same amide isomer. The right panel compares the IRMPD spectrum with the predicted spectrum for the lowest-energy

conformer of the enol-imine isomer, which lies 60 kJ/mol above the lowest-energy amide conformer.

In the calculated enol-imine spectrum, two relatively strong diagnostic bands (*) are due to NH bending of the enol-imine moiety (1100 cm^{-1}) and to a delocalized bending vibration involving enol-imine NH and OH bonds (840 cm^{-1}). In the computed spectrum for the amide isomer, no strong bands are predicted at these frequencies. The observation that the experimental spectrum shows little or no intensity at these frequencies suggests that the c_3-ion possesses an amide structure. The remainder of the experimental spectrum between 1200 and 1750 cm^{-1} also shows good agreement with the spectrum predicted for the amide isomer, especially for the conformer at slightly elevated energy (blue trace).

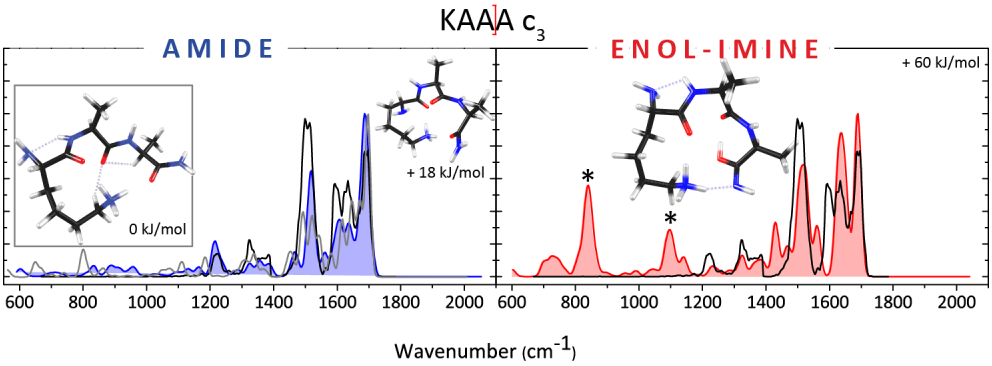


Figure 8-1. *Experimental infrared spectrum of the ETD generated c_3-ion from $[KAAA+2H]^{2+}$ (in black) compared with computed spectra for the lowest energy amide structure (left, gray), a higher energy conformation of the amide isomer (left, blue), and the lowest energy enol-imine structure (red, right). The experimental spectrum is assigned as an amide structure based on the generally favorable overlap between the experimental and computed spectrum in blue and in particular on the absence of O-H and N-H bending modes in the experimental spectra, diagnostic for the enol-imine structure (indicated with asterisks on the right).*

This isomeric structure assignment is in line with that for two *c*-type ExD fragments studied by IR spectroscopy previously.[64,65] Ref.[64] addresses the structure of a (modified) c_0-ion, so that the two isomeric product ions can conveniently be distinguished based on the absence or presence of the amide C=O stretch (at 1731 cm^{-1}). The enol-imine structure does not possess a carbonyl moiety so that its predicted spectrum does not contain a C=O stretch band. The longer c_3 and c_4 ions studied in the present work possess multiple amide C=O moieties – one for each

peptide linkage – so that the presence or absence of a carbonyl stretch is no longer a good diagnostic. Instead, the presence of the NH bending mode near 1100 cm^{-1} is used here as a diagnostic for the enol-imine structure.[65]

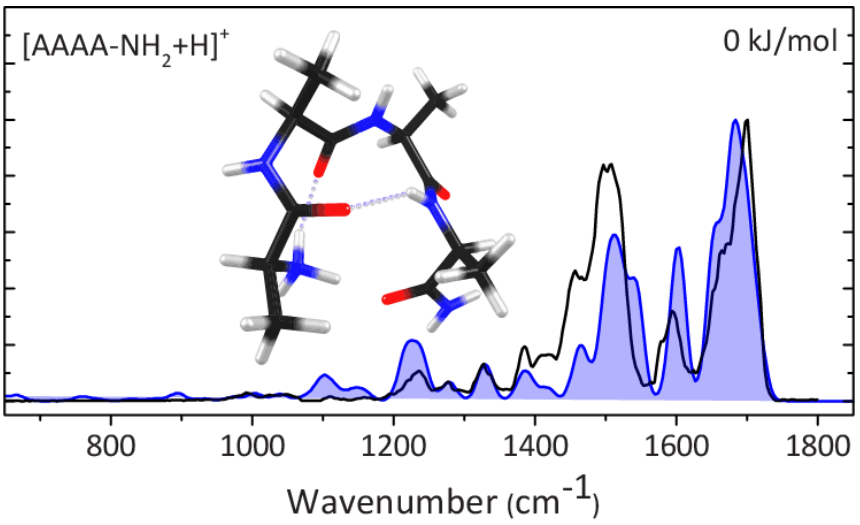


Figure 8-2. *Experimental spectrum of NH_2-capped AAAA compared with the computed spectrum for a N-terminally protonated structure.*

Peptides with a C-terminal NH_2-cap are readily available and can serve as a reference for *c*-type fragment ions in the amide isomeric configuration.[64,65,67,72] The IRMPD spectrum of protonated NH_2-capped tetra-alanine, $[AAAA\text{-}NH_2+H]^+$, is shown in Figure 8-2. The band near 1700 cm^{-1} is due to C=O stretches of the first and last peptide linkages, with the shoulder at 1670 cm^{-1} being due to a combined C=O stretch with N-H bending at the N-terminus of the peptide. The peak at 1600 cm^{-1} corresponds to NH bending at the protonated N-terminus and that around 1500 cm^{-1} to modes with combined backbone NH and terminal NH_3 bending character. Clearly, matching this feature with theory appears slightly more challenging than the other bands, which we shall keep in mind in our analysis below.

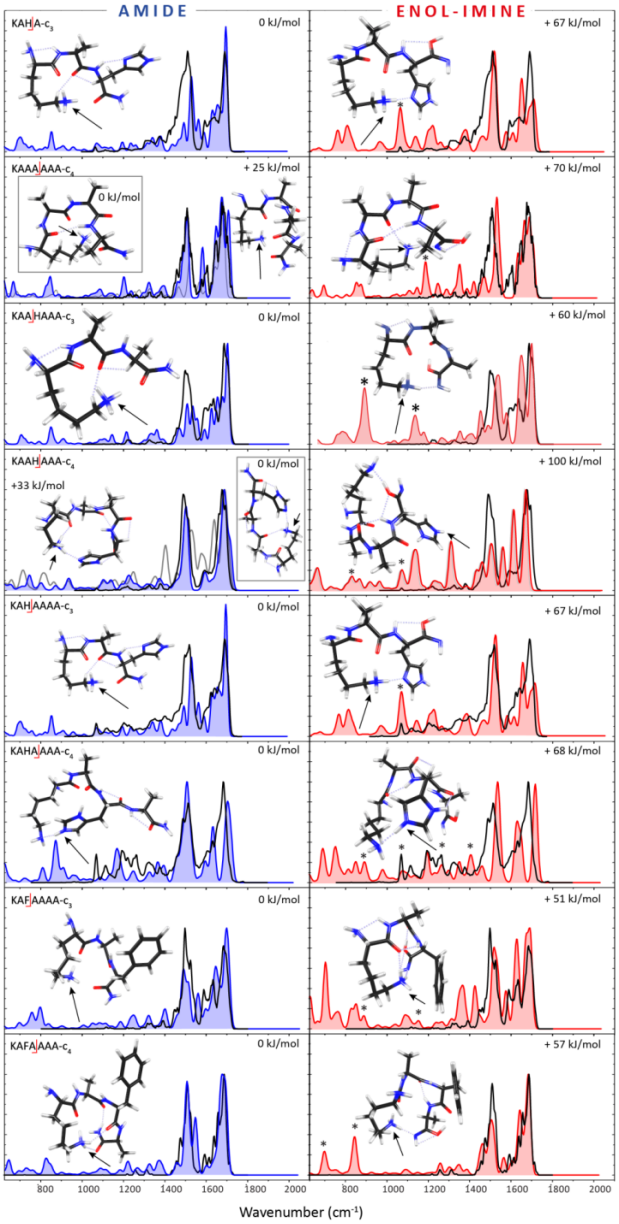


Figure 8-3. *Experimental spectra of a series of c_3 and c_4 ETD fragment ions (black) compared with the computed spectra for the amide structures (left, blue) and for the enol-imine isomers (red, right). The stars indicate the enol-imine NH and OH bending modes. Arrows indicate the protonation site. For the c_4-ions of KAAAAAA and KAAHAAA, the best match is found for a higher-energy conformer; the spectrum of the lowest-energy amide conformer is shown in addition in light grey in these cases.*

Relying on the good overall match between experiment and theory in Figure 8-2 and having established the salient distinguishing features between amide and enol-imine IR spectra, we extend our study to a large set of c_3 and c_4 ions generated from different precursor peptides. Figure 8-3 shows the experimental IRMPD spectra of eight different c_3 and c_4 ions obtained from ETD on doubly protonated KAHA, KAAAAAA, KAAHAAA, KAHAAAA and KAFAAAA. All experimental spectra are compared with the calculated spectra for both the enol-imine and the amide isomer, including different conformations of these isomers. In Figure 8-3, the left panels compare experimental spectra with computed spectra for amide isomers, with the best matching spectra in shaded blue and, if different, the most stable conformer in gray. The panels on the right show the predicted spectra for the alternative enol-imine isomers of each of the *c*-ions overlaid on the same experimental spectra; diagnostic NH and OH bending modes are indicated by asterisks. The spectrum for the lowest-energy conformer is shown in all cases.

The c_3-ions of KAHAAAA and of KAHA likely are identical, which is indeed confirmed by their IR spectra, see Figure 8-4. The computed spectra on the first and fifth row of Figure 8-3 are therefore identical. Similarly, the c_3-ions of KAAA and KAAHAAA are also the same as suggested by an overlay of their IR spectra in Figure 8-4; the computed spectra in Figure 8-1 and on the third row of Figure 8-3 are therefore identical.

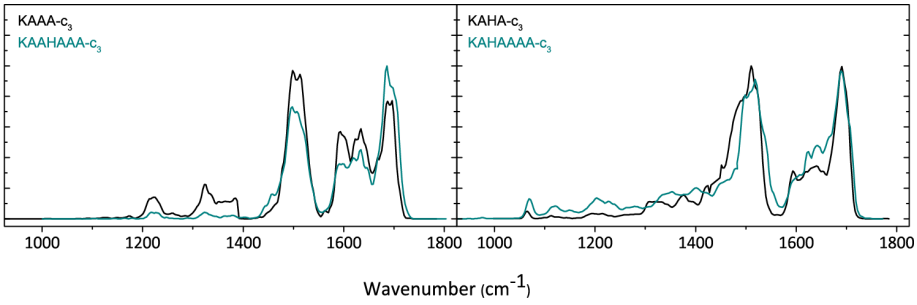


Figure 8-4. *Overlay of the experimental IRMPD spectra of the c_3-ions of doubly protonated KAAA and KAAHAA (left panel), and between the c_3-ions of KAHA and KAHAAAA (right panel). Ignoring a small apparent frequency miscalibration between 1300 and 1600 cm^{-1} in the right panel, the two spectra are nearly identical, which we take as evidence for the two pairs of c_3-ions to possess identical structures. We conclude that the different neutral loss fragments do not influence the structure of the c-ions formed.*

For all structures, the preferred protonation site is the Lys side chain, except for the c_4-ions of KAAHAAA and KAHAAAA, where protonation occurs on the His residue for the enol-imine isomers. The amide structure of KAHAAAA-c_4 is also protonated at the His residue. Some enol-imine conformers converge to a structure where the proton has transferred from the Lys or His side chain to the imine nitrogen. As a consequence, their calculated spectra do not possess the typical NH bending modes. The match with the experimental spectra is poor as shown in Figure 8-5 and we discard these structures.

All spectra feature prominent amide I (backbone amide carbonyl stretching around 1600 – 1700 cm^{-1}) and amide II (backbone amide NH bending around 1500 cm^{-1}) bands. The spectra in the 1400 – 1700 cm^{-1} range are generally well reproduced by the computed spectra for the amide isomers in blue except for some deviations in intensity for amide II in a few cases, as already noted for the NH_2-capped Ala_4 reference (Figure 8-2). As for the c_3 ion of KAAA in Figure 8-1, the lowest-energy conformer does not always provide the best match; for the c_4-ions of KAAAAAA and KAAHAAA the spectral match is better for higher-energy conformers, which we tentatively attribute to kinetic trapping.

As compared to the amide spectra, the calculated spectra for the enol-imine isomers qualitatively display more prominent deviations from experiment in the 1400 – 1700 cm^{-1} range, with for instance significant mismatches in the amide I band (*e.g.* KAH-c_3, KAA-c_3 and KAAH-c_4) and the absence of the 1590 cm^{-1} band for KAAA-c_4. Also at lower frequencies, amide spectra appear to provide an overall closer agreement than enol-imine spectra, with the general absence of the enol-imine NH and OH bending modes (*) in the experimental spectra as specific examples. The overall picture that emerges from the comparisons in Figure 8-3 is that the amide isomer spectra provide the best match to the experimental spectra.

The enol-imine structures of interest here are between 51 and 100 kJ/mol higher in energy than the global minimum amide structures, which is in line with relative energies reported previously[64,66,68,73]. Although the IR spectra suggest an amide structure for all *c*-ions studied here, we cannot exclude the possibility that the amide structure is formed via an enol-imine intermediate, driven by the substantial gain in energy. The energetic barriers involved in such transformations have been investigated by transition-state (TS) calculations and it appears likely that isomerization occurs before the incipient *c*- and *z*-fragments separate.[68,73] The charge-solvating behavior of the *z*-fragment in this pre-dissociation ion-molecule

complex lowers the barriers for the H-atom transfer and can be regarded as a catalyst for the isomerization.[73] The actual TS energies depend on the specific system, but are in all cases much lower than the energy available from the charge recombination process.[74]

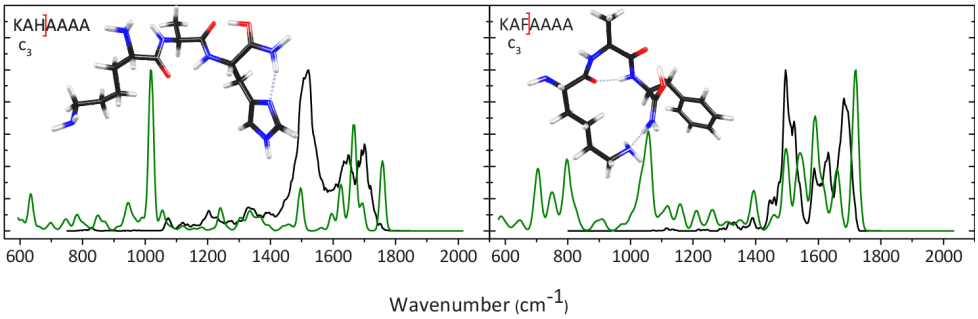


Figure 8-5. *The experimental spectra of the c_3-ions of KAHAAAA and KAFAAAA compared with calculated spectra of structures that have formed by proton transfer from the Lys or His side chain to the enol-imine nitrogen atom. The calculated spectra for these tautomers do not possess the typical NH-bending mode around 1100 cm^{-1}. Overall, we conclude that the match with the experimental spectra is poor and we discard these structures.*

8.4. Conclusion

In conclusion, based on both our structural assignments from IR spectral matching in Figures 8-1 and 8-3 and the corresponding computed relative energies, this study suggests that for a set of nine *c*-type ions, all most likely possess an amide C-terminus. Assuming the set of sequences selected for the precursor peptides is sufficiently diverse to be generally representative for a broader range of peptides, these results suggest *c*-type ions not having C-terminal amide groups would be unusual and that they generally are NH_2-capped truncated peptides. One caveat may be the presence of a Lys residue in all systems studied here. Based on computational investigations, the ammonium group of a protonated Lys side chain was suggested to catalyze the enol-imine to amide tautomerization.[68] However, isolated cases of *c*-ions not including a Lys residue have been spectroscopically investigated and gave evidence for amide structures without exceptions.[64,65,75] The bottom line here is that the enol-imine structure was not encountered in any of the spectroscopic investigations thus far. These experiments do not exclude the possibility that enol-imine structures are traversed as reaction intermediates, but may be regarded as experimental evidence for their efficient conversion to amide structures.[68,73]

Chapter 9

Investigation of the position of the radical in z_3-ions resulting from electron transfer dissociation using infrared ion spectroscopy

---

The molecular structures of six open-shell z_3-ions resulting from electron transfer dissociation mass spectrometry (ETD MS) were investigated using infrared ion spectroscopy in the 800-1850 and 3200-3700 cm^{-1} spectral range in combination with density functional theory and molecular mechanics/molecular dynamics calculations. We assess in particular the question of whether the radical remains at the C_α-site of the backbone cleavage, or whether it migrates by H-atom transfer to another, energetically more favorable position. Calculations performed herein as well as by others show that radical migration to an amino acid side chain or to an α-carbon along the peptide backbone can lead to structures that more stable, up to 33 kJ/mol for the systems investigated here, by virtue of resonance stabilization of the radical in these alternative positions. Nonetheless, for four out of the six z_3-ions considered here, our results quite clearly indicate that radical migration does not occur, suggesting that the radical is kinetically trapped at the site of ETD cleavage. For the two remaining systems, a structural assignment is less secure and we suggest that a mixture of migrated and unmigrated structures may be formed.

* Adapted from: Kempkes, L.J.M.; Martens, J.; Berden, G.; Houthuijs K.J.; Oomens, J.; Investigation of the position of the radical in z3-ions resulting from electron transfer dissociation using Infrared Ion Spectroscopy. *Faraday Discuss., In Press*

9.1. Introduction

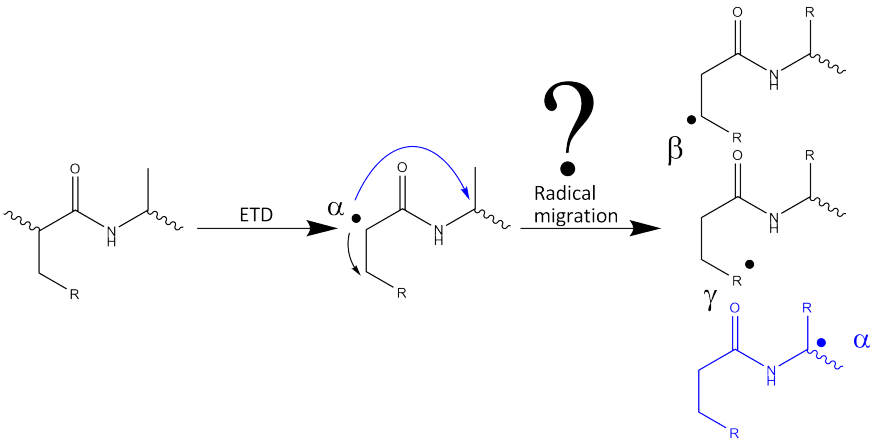
Electron Capture Dissociation (ECD) and Electron Transfer Dissociation (ETD) have developed into popular MS/MS methods in top-down proteomics. Advantages of ExD dissociation methods over collision-induced dissociation include an overall improved sequence coverage and retention of labile post-translational modifications (PTMs). However, ExD mass spectra are more complex and are often more difficult to interpret, especially because the radical-driven reaction mechanisms remain incompletely understood at a fundamental level.

ExD commonly induces cleavages of the protein backbone at the N-C_α bonds, leading to N-terminal c-type fragment ions and radical C-terminal z-type sequence ions.[24,28-35] The reaction mechanisms leading to the backbone cleavage and the resulting fragment structures are the subject of extensive discussion in the scientific literature[17,24,27,29,30,38-51,76]. The position of the radical in the peptide and its ability to migrate along the peptide backbone or to the amino acid side chains in the open-shell fragment ions have been explored in a large number of experimental and theoretical investigations.[57,66,67,75,77-99]

The so-called Cornell[17,29,39,53] and Utah-Washington[34,38,40,44-47,52,54,56,57] mechanisms are the best-known hypotheses describing ExD dissociation reactions. In the Cornell mechanism, electron attachment occurs at a protonated site in the ion that is hydrogen-bonded to a nearby carbonyl, whereas according to the Utah-Washington mechanism, the electron is directly captured in a π^* orbital of an amide carbonyl group, stabilized by H-bonding to a protonated site. In both mechanisms, this eventually leads to N - C_α bond cleavage, although the resulting c-type fragment structures may be different. In contrast, both mechanisms agree on the C-terminal z-type ions being formed with the radical initially at the C_α-position of the cleaved peptide bond, as shown in Scheme 9-1. However, this is in most cases not the thermochemically most favored structure and radical migration to form lower-energy isomeric species has been suggested to occur.[100] Radical migration may lead to various secondary side chain losses and backbone dissociations at sites distant from the location of the initial radical.[100] Radical migration in z-ions has therefore been investigated by CID MS/MS,[26,67,77-79,84,99,101] and also by UV photodissociation spectroscopy (UV-PD),[57,80-82,89,102,103] ion mobility[66] and infrared ion spectroscopy[75].

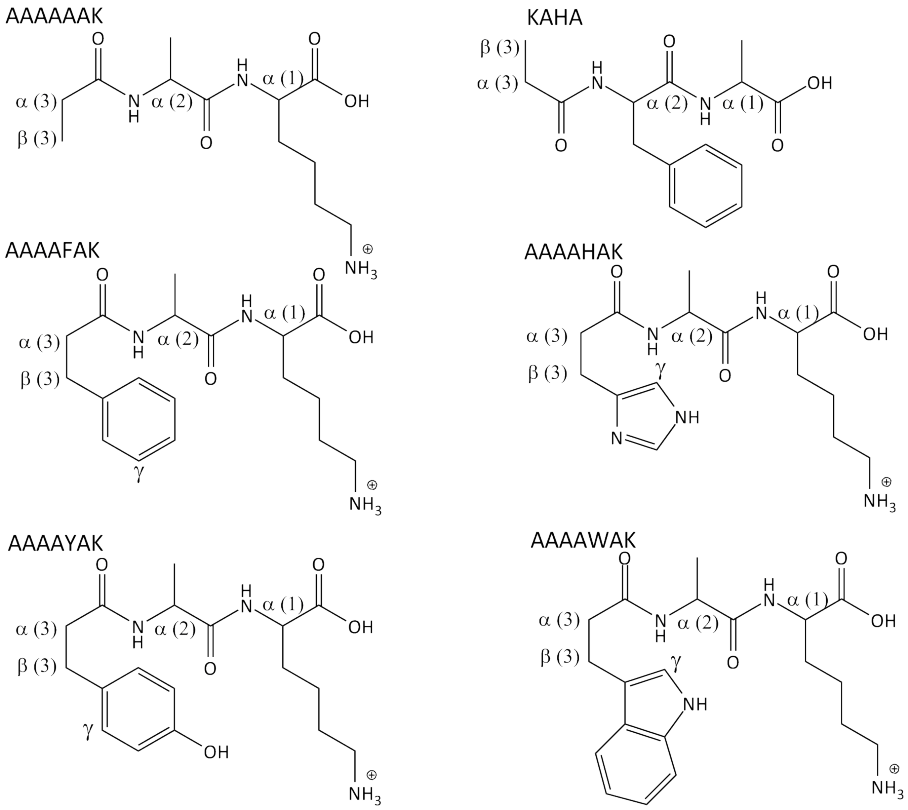
CID methods have been employed to provide information about radical migration by activation of the z-ion and analysis of the resulting CID reaction products.[104-106] The location of the unpaired electron in the ExD-generated z-type radical cation is

assumed to be a major factor determining the cleavages of the backbone or the side chains of the peptide.[107] For example, CID has been applied on a systematically varied set of z-ions resulting from ETD on the pentapeptides $[AAXAR+2H]^{2+}$.[77,99] For z-ions resulting from peptides with aromatic residues (X = F, H, Y and W), very facile radical-induced backbone cleavages were observed, which was attributed to the readily transferrable side-chain β-hydrogen atom,[77,99] allowing for resonantly-stabilized β-radical structures to be formed. A caveat in CID MS^n based studies is the possibility that the radical migrates due to the collisional activation used to probe the ExD reaction, but not due to the ExD reaction itself.[92]



Scheme 9-1. *Schematic representation of the formation of z-type product ions formed by Electron Transfer Dissociation and radical migration within the product ion. The radical is able to migrate within the amino acid sidechain from the cleavage site (C_α-position) to the adjacent methyl groups (C_β- (top) and C_γ- (middle) positions). Besides radical migration within the amino acid residue, the radical can also migrate to adjacent residue(s), represented by the blue arrows leading to the bottom α-radical structure.*

Comparison of calculated collisional cross sections (Ω) with cross sections determined by ion mobility has been used to investigate the structures of c- and z-ions resulting from ETD on $[AAHAL+2H]^{2+}$.[66] For the z_3 fragment, the position of the radical has been assigned as the C_α of the His residue at the cleavage site, indicating no radical migration.[108] For the z_4 ion, the Ω-values were too similar to distinguish between the different radical positions. Based on ion-mobility data for the charge-reduced but undissociated ion (ETnoD), which showed that hydrogen atom migrations were slow in comparison with the drift time, the location of the radical in the z_4-ion was derived to be the original $C_\alpha(4)$ position at the ETD cleavage site.

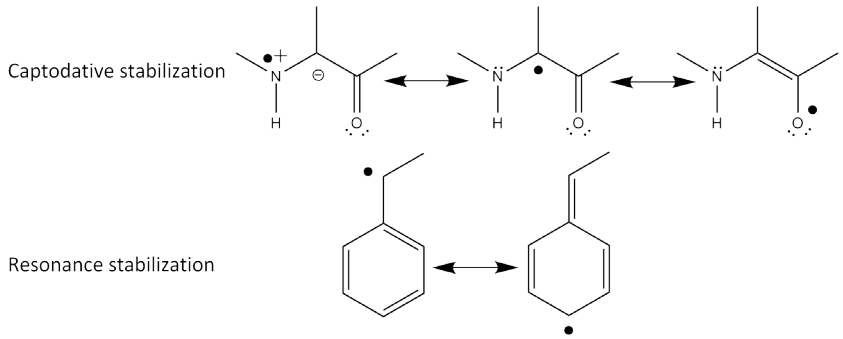


Scheme 9-2. *Structures of the investigated z_3-ions, with the possible α, β and γ positions of radical indicated for each z-ion.*

Infrared ion spectroscopy was applied to a series of radical z-ions (z_1-, z_2-, z_3- and z_4-ion) generated by ETD from doubly charged $[AAHAR+2H]^{2+}$. In all z-ions the radical was established to be located at the α-carbon of the cleavage site, suggesting that radical migration to thermodynamically more favorable sites does not occur.[75] Tureček and coworkers extensively investigated radical migration in z-type ions using UV photodissociation (UVPD) spectroscopy. For the z_4 ions of $[ADAAR+2H]^{2+}$ and $[QDAAR+2H]^{2+}$, the UVPD action spectrum was attributed to a mixture of radical isomers by H-atom migration over the backbone or to the aspartic acid C_β-position. Also for z_4 ions $[^{\bullet}NAAR+H]^{+}$, $[^{\bullet}EAAR+H]^{+}$ and $[^{\bullet}QAAR+H]^{+}$, the UVPD action spectra indicate that partial isomerization occurred upon ETD.[57] The UVPD spectra of z_4 ions formed by ETD on $[AAWAR+2H]^{2+}$ and $[AAFAR+2H]^{2+}$ were attributed to structures with the radical at the cleavage site, although the presence

of other isomers could not be excluded[80]. Calculations indicated that radical migration from the initial cleavage site to the C_α of Arg leads to the most stable isomer for the z_4 ion of $[AAWAR+2H]^{2+}$.[80] For the z_4 ion of $[AAFAR+2H]^{2+}$, calculations suggest that radical migration from the Ala residue at the cleavage site to the C_β of Phe gives the most stable isomer. Nonetheless, the UVPD spectrum for this ion suggests an unmigrated radical structure, although this assignment was less secure.

Here we use IRMPD ion spectroscopy to explore the influence of aromatic residues at the cleavage site, with Ala serving as a non-aromatic reference. We focus on the question of whether or not radical migration occurs in the series of investigated z_3-ions, which are derived from applying ETD on heptapeptides containing an aromatic residue in the third position (counted from the C-terminus). In addition, the z_3 ion from $[KAHA+2H]^{2+}$ has also been investigated. Scheme 9-2 lists the structures of the z_3 ions that have been investigated, with possible positions of the radical (α-,β-, and γ-carbons) indicated.



Scheme 9-3. *Resonance stabilization of the radical positioned at the C_β position in the aromatic residues (bottom) and captodative stabilization of the radical at the C_α position of the first or second residue (top).*

Of these possible positions of the radical, the $C_\alpha(2)$- and $C_\alpha(1)$-positions in the protein backbone and the $C_\beta(3)$-position in the aromatic residues typically generate structures having the lowest energy. The radical at the C_α position in the neighboring residues is stabilized by the captodative effect,[109-113] which can be characterized as π-electron delocalization between the adjacent electron-withdrawing carbonyl group and the electron-donating amide nitrogen. This effect has for instance been extensively described and studied using ion spectroscopy for the histidine amino acid radical [98] and is illustrated in the top panel of Scheme 9-3. Radical migration to the C_β position within the aromatic residue is stabilized by

delocalization of the electron over the aromatic system (resonance stabilization), as illustrated in the bottom panel of Scheme 9-3.[114] Finally, the $C_\alpha(3)$-position at the cleavage site is of particular interest here. Although this isomer is usually relatively high in energy, $C_\alpha(3)$ is the initial site of the radical according to the Cornell and Utah-Washington mechanisms. If radical migration faces high barriers, the system may remain kinetically trapped in this isomer.

9.2. Experimental and computational methods

9.2.1. IRMPD spectroscopy

An ETD-enabled 3D quadrupole ion trap mass spectrometer (Bruker, AmaZon Speed ETD), modified to enable optical access to the ion population, has been used to produce the z_3-ions and to obtain their infrared spectra.[37,58,75,108,115] Precursor peptide samples were purchased from Biomatik (Canada) and were used without further purification. Doubly protonated precursor ions were generated by electrospray ionization (ESI) from ~10^{-6} M solutions in 50:50 acetonitrile/water with ~0.5 % formic acid added. The doubly charged precursor ion of interest was mass isolated and stored in the quadrupole ion trap. Fluoranthene radical anions generated in the chemical ionization source of the instrument were then guided into the trap, where the ETD reaction with the peptide dications took place for 300 ms. The fragment ion of interest is then mass-isolated for spectroscopic interrogation.

The mass spectrometer is coupled to the optical beam line of the free electron laser FELIX[116], which enables us to record infrared multiple-photon dissociation (IRMPD) spectra in the 800-1850 cm^{-1} region; an optical parametric oscillator (OPO, Laser Vision, Bellevue, USA) is used for measurements in the 3200-3700 cm^{-1} region [58]. For the FELIX experiments, the z-type fragment ions were irradiated with two IR pulses at a repetition rate of 10 Hz at pulse energies of approximately 20 – 60 mJ. The bandwidth is around 0.5% of the center frequency. The OPO generates pulses of approximately 15 mJ at a 10 Hz repetition rate with a bandwidth of 3 cm^{-1}. The dissociation yield at each IR frequency is obtained from five averaged mass spectra and calculated as ΣI(fragment ions)/ΣI(parent + fragment ions). By plotting the yield as a function of laser frequency, an infrared spectrum is generated. No apparent differences were noticed between the wavelength dependency of fragmentation into the individual channels; no clear evidence was therefore found for the co-existence of multiple isomers, such as for instance in some of our previous CID MS/MS studies.[59,117] The yield is linearly corrected for the frequency-dependent

change in laser pulse energy and the infrared frequency is calibrated using a grating spectrometer (FELIX) or a wavemeter (OPO).

9.2.2. Computational chemistry

A Molecular Mechanics/Molecular Dynamics (MM/MD) approach was applied to z_3 ions with the radical positioned at the cleavage site $C_\alpha(3)$, at the β- and γ- positions within the terminal residue ($C_\beta(3)$ and $C_\gamma(3)$) and at the α-position of the second residue (alanine, $C_\alpha(2)$). AMBER 12 [70] was used to explore the potential energy surface of the z-ions and to find the lowest-energy conformers. After minimization within AMBER, a simulated annealing procedure up to 500 K was used. 500 structures were obtained in this procedure and grouped based on structural similarity using appropriate rms criteria to give 20 to 30 candidate structures. Next, these structures were each optimized using Density Functional Theory (DFT) at the UB3LYP/6-31++G(d,p), the UB3LYP-D/6-31++G(d,p) and the UM06-2X/6-31++G(d,p) levels of theory. IR spectra were also predicted at these levels of theory and we use the B3LYP spectra for comparison with experiment.

All DFT calculations are performed using the Gaussian09 revision D01 package[69]. The computed harmonic vibrational frequencies were scaled by 0.975 in the 800-1850 cm^{-1} region and by 0.955 for the 3200-3700 cm^{-1} region. To facilitate comparison between calculated and experimental spectra, the calculated spectra were convoluted with a 25 cm^{-1} full-width-at-half-maximum Gaussian line shape. The computational procedure is described in more detail elsewhere[71] and has for instance applied in our studies on ETD generated w-ions and c-ions.[108,115].

9.2.3. Spectral matching scores

To assess the degree of similarity between calculated and experimental infrared spectra, we employ the cosine similarity as an objective measure.[118] The cosine of the angle θ between two *n*-dimensional vectors is calculated using their normalized Euclidean dot product according to

$$similarity = \cos(\theta) = \frac{\boldsymbol{A}\cdot\boldsymbol{B}}{\|\boldsymbol{A}\|\|\boldsymbol{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \qquad [9\text{-}1]$$

where **A** and **B** are two *n*-dimensional vectors with A_i and B_i as their *i*-th elements. This method assesses the degree to which the two vectors, representing the experimental and theoretical spectra in this case, are parallel. A cosine value closer to unity indicates that experimental and computed spectra are more similar. The intensity values in the computed spectrum are evaluated at the exact

wavenumbers of the experimental spectrum, so that the two spectra have a common *x*-axis.

We slightly modify the above equation to make the cosine similarity index less sensitive to deviations in peak intensities between ***A*** and ***B***, and thereby more sensitive to frequency overlap between bands in ***A*** and ***B***. Both the experimental and calculated spectra are scaled to a maximum intensity of 1 and then the logarithm of these scaled values is taken

$$A_i^{trans} = \log(\frac{A_i}{A_{max}} + c) \qquad [9\text{-}2]$$

where *c* is a constant that is identical for vectors ***A*** and ***B***. The value of *c* is a compromise between being sensitive to low-intensity bands in the spectrum on the one hand and avoiding experimental noise to affect the similarity on the other hand. A value 10^{-8} was taken here for *c* as it gave the best results for small test set of experimental and computational spectra. The cosine similarity is calculated for the spectra and displayed in the figures as the 'matching score'.

9.3. Results and discussion

For each of the z_3 ions studied here, Table 9-1 lists the relative energies for isomers differing in the position of the radical. We adopt here a numbering of the residues and molecular groups that starts from the C-terminus. For each isomer, the B3LYP/6-31++G(d,p) relative free energy of the lowest-energy conformer identified in the conformational search is presented. For structures with the radical on the C_β atom of the second residue and on the C_α atom of the first residue, no conformational search was performed as these structures are unlikely (see below); Table 9-1 lists the DFT energy that may not correspond to the lowest-energy conformer. A complete overview of the relative energies including results obtained at B3LYP-D and M06-2X levels of theory is presented in Supplementary Information Table S1 in the appendix of reference[119].

The values reported in Table 9-1 indicate that isomers with the radical positioned at the cleavage site, $C_\alpha(3)$, are relatively high in energy. Radical migration to the β-position of the same residue leads to significantly lower-energy isomers for aromatic residues, due to resonance stabilization of the radical by conjugation with the aromatic moiety. Radical migration to the α-positions along the backbone (1st and 2nd residue from C-terminus), leads to isomers that are 15 to 32 kJ/mol more stable as a consequence of captodative stabilization of the radical.

		3rd residue		2nd residue		1st residue
Peptide|fragment	C_α	C_β	C_γ	C_α	C_β	C_α
AAAA|AAK	**+ 21**	+ 63	-	**0**	+ 71	+ 1*
AAAA|FAK	+ 32	+ 3	+ 102	**0**	+ 64	+ 8*
AAAA|HAK	+ 22	+ 17	+ 153	+ 8	+ 68*	**0***
AAAA|WAK	+ 31	+ 2	-	**0**	+ 72*	+ 56*
AAAA|YAK	+ 33	**0**	+ 111	+ 15	+ 77*	+ 7*
K|AHA	+ 15	+ 66*	-	**0***	+ 11*	+ 17*

Table 9-1. *Computed relative energies for the various possible isomers of the radical z-ions investigated at the B3LYP/6-31++G(d,p) level of theory. Residues are counted from the C-terminus. * No MM/MD conformational search was performed for these structures (see text).*

An extensive conformational search was not deemed necessary for isomers with the radical at the C_β(2) position, because this position does not enable resonance stabilization of the radical and gives high-energy structures. This is specifically verified for z_3-ions of AAAAAAK and AAAAFAK, which gave significant spectral mismatches around 1500 – 1700 cm^{-1}, as shown in Figure 9-1.

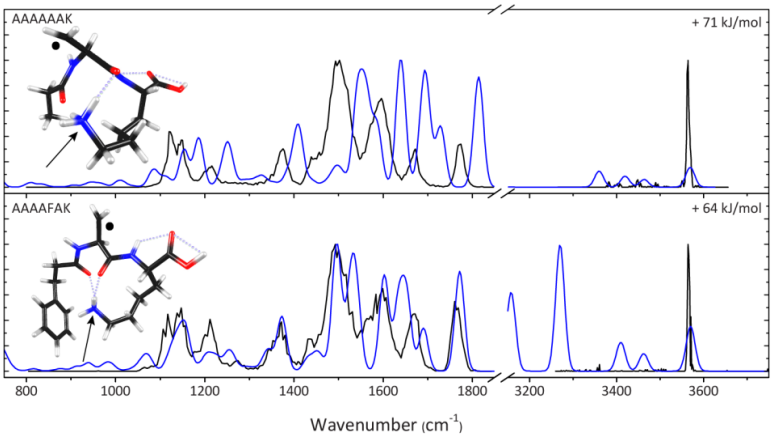


Figure 9-1. *Experimental spectra of the z_3-ions from $[AAAAAAK+2H]^{2+}$ and $[AAAAHAK]^{2+}$(black) compared with calculated spectra (blue) for structures with the radical at the C_β position of the second residue (Ala). These structures are relatively high in energy due to the absence of resonance stabilization of the radical.*

Figure 9-2 compares experiment with calculated spectra for structures in which the radical has migrated to the $C_\alpha(1)$ position. All experimental IR spectra show a distinct band around 1750-1770 cm^{-1}, which is attributed to the C=O stretch of the carboxylic acid group at the C-terminus. In the computed spectra, this C=O stretch undergoes a significant red shift as a consequence of conjugation with the radical at the adjacent α-carbon, effectively reducing the C=O bond order.[75] Hence, the $C_\alpha(1)$ radicals are clearly unable to explain the experimental spectra and we can safely exclude these structures, even if they are low in energy; no conformational search has been carried out for these isomers.

Figure 9-3 presents the relative energies, matching scores, fragmentation channels, experimental and calculated spectra for the series of z_3-ions $[^\bullet AAK+H]^+$, $[^\bullet HAK+H]^+$, $[^\bullet YAK+H]^+$ and $[^\bullet AHA+H]^+$. The radical positions that are considered here are: the cleavage site, $C_\alpha(3)$, the β-carbon of the third residue, $C_\beta(3)$, and the α-carbon of the adjacent residue, $C_\alpha(2)$. For the z-ions with an alanine residue in the third position ($[^\bullet AAK+H]^+$ and $[^\bullet AHA+H]^+$), radical migration to the $C_\beta(3)$-position does not enjoy resonance stabilization and leads to relatively high-energy structures; we therefore consider these isomers unlikely and their calculated spectra are not included in Figure 9-3.

For the $[^\bullet AAK+H]^+$ z_3-ion, the lowest-energy structure is found for the $C_\alpha(2)$ isomer (21 kJ/mol more stable than the $C_\alpha(3)$ cleavage-site radical). A gain of 14 kJ/mol in energy upon radical migration to the $C_\alpha(2)$ position is observed for the $[^\bullet HAK+2H]^+$ z_3-ion, while migration to the $C_\beta(3)$-position of the His residue gives only a marginal stabilization (-5 kJ/mol). For the $[^\bullet YAK+2H]^{2+}$ z_3-ion, the lowest-energy isomer is that with the radical at the $C_\beta(3)$-position of Tyr. For the $[^\bullet AHA+2H]^{2+}$ z_3-ion, radical migration to the $C_\alpha(2)$-position of His stabilizes the system by 15 kJ/mol. Hence, based purely on thermochemical grounds, radical migration from the original $C_\alpha(3)$-position of the ETD cleavage site to the $C_\beta(3)$-position or along the peptide backbone to adjacent C_α-positions is plausible.

In addition to the relative energies of the isomers (see above) and the IR spectral information (see below), the observed IR-induced fragmentation channels may also contain information on the structure of the z_3-ions. However, one should keep in mind that IR multiple-photon excitation is a slow process and that rearrangements such as H-atom migration may occur upon activation prior to dissociation. A similar reservation was made concerning CID probing of radical migration.[92]

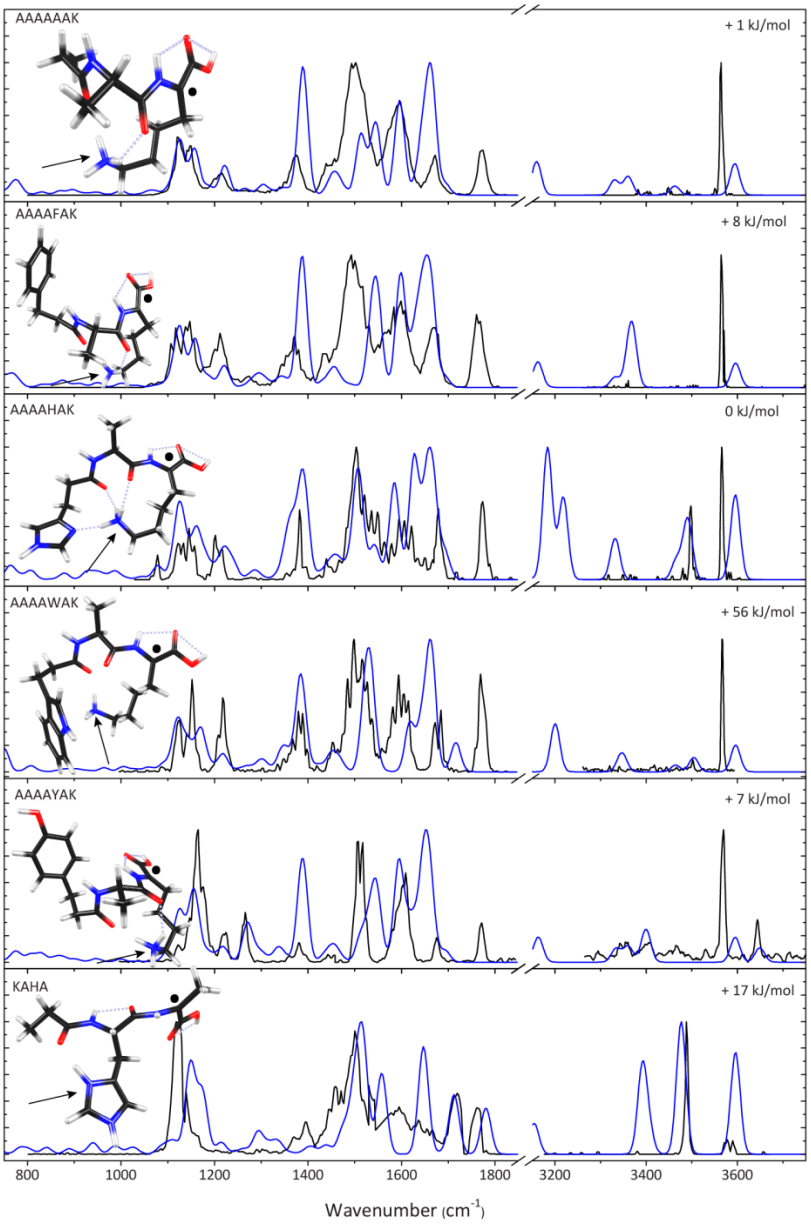


Figure 9-2. *Experimental IRMPD spectra (in black) of the six z_3 ions considered in this study compared to calculated spectra (blue) for z_3 ions with the radical at the C_α of the first residue (lysine, or alanine for KAHA). A striking mismatch between experiment and theory is observed for the C-terminal carboxylic C=O stretch mode in the 1700 – 1800 cm^{-1} range; $C_\alpha(1)$-radical structures are therefore rejected.*

For the systems in Figure 9-3, no radical side-chain loss is observed for $[^{\bullet}AAK+H]^+$, $[^{\bullet}AHA+H]^+$ and $[^{\bullet}YAK+H]^+$ suggesting that the radical does not migrate to the side chains. For $[^{\bullet}HAK+H]^+$ and $[^{\bullet}YAK+H]^+$ the formation of a z_2-ion is observed, suggesting that the radical is able to migrate to the $C_\alpha(2)$ position; as mentioned, we cannot exclude that this happens only upon IR activation.

We now inspect the IR spectra in Figure 9-3. For systems with an aromatic residue in position 3, a $C_\beta(3)$ radical z_3-ion structure is conceivable and their calculated spectra are shown in green. Both computed spectra predict an intense peak around 1650 cm^{-1} (C=O stretching in the Tyr or His amino acid residue, combined with NH_3 bending) and around 1550 cm^{-1} (NH_3 bending). Both of these features are poorly reproduced in the experimental spectrum. Based on this severe mismatch, the resonance-stabilized $C_\beta(3)$ structure is rejected for the $[^{\bullet}YAK+H]^+$ and $[^{\bullet}HAK+H]^+$ z_3-ions.

Next we consider the $C_\alpha(2)$ radical structures that are stabilized by the captodative effect, whose calculated spectra are shown in red in Figure 9-3. Again, the spectral region between 1650 and 1550 cm^{-1} containing the C=O stretches is poorly reproduced by the calculations. Also, for $[^{\bullet}AHA+H]^+$ the doublet feature between 1700 cm^{-1} and 1800 cm^{-1} is not correctly reproduced by the calculation. Finally, for all systems in Figure 9-3 except $[^{\bullet}AHA+H]^+$, the bands near 1400 cm^{-1} are slightly red-shifted compared to the experimental spectra. Based on these observations the possibility of radical migration to the C_α-position of a neighboring residue is unlikely.

Computed spectra for structures with the radical positioned at $C_\alpha(3)$, the ETD cleavage site, are colored blue in Figure 9-3. Of the three types of isomers considered in this Figure, these structures provide the best match for the experimental feature around 1675 cm^{-1}, especially for $[^{\bullet}HAK+H]^+$, assigning this peak as an amide C=O stretch. Overall, the calculated spectra for the $C_\alpha(3)$ structures that have not undergone radical migration show the best agreement in the amide 1400 – 1650 cm^{-1} region. Especially in the experimental spectra of $[^{\bullet}AAK+H]^+$ and $[^{\bullet}YAK+H]^+$, the C=O stretch band of the second amide linkage around 1600 cm^{-1} is reproduced better than by the alternative structures. It is also striking that the predicted spectra for the $C_\alpha(3)$ radical reproduce the experimental spectra accurately in the entire 1000 – 1400 cm^{-1} region, whereas the alternative $C_\beta(3)$ and $C_\alpha(2)$ spectra fail to do so.

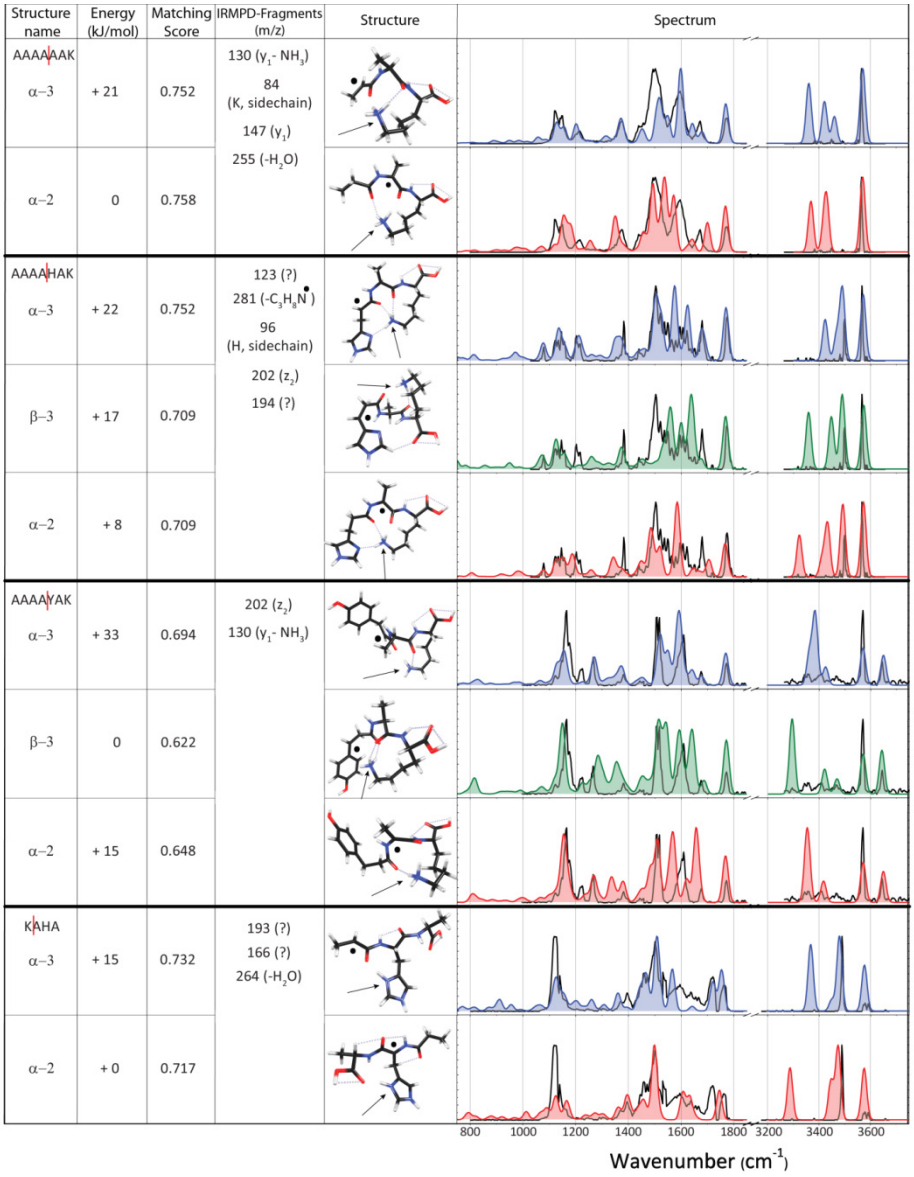


Figure 9-3. *IRMPD spectra (black) of the four z_3 ions $[^{\bullet}AAK+H]^+$, $[^{\bullet}HAK+H]^+$, $[^{\bullet}YAK+H]^+$ and $[^{\bullet}AHA+H]^+$ compared to calculated spectra with the radical positioned at $C_\alpha(3)$ in blue, at $C_\beta(3)$ in green and at $C_\alpha(2)$ in red. The presented calculated spectra are for the lowest energy conformers for each z_3-isomer resulting from the MM/MD search.*

The $[^{\bullet}AHA+H]^+$ z_3-ion is an exception in the present series of investigated systems in that the $C_\beta(2)$-position is of interest for radical migration, due to stabilization through resonance with the imidazole group of the His residue. The computed spectrum for this isomer is overlaid onto the experimental spectrum in Fig. S2 in ref. [119], and the agreement between computed and experimental spectra is seen to be comparable to that for the $C_\alpha(3)$ isomer. Radical migration to the $C_\beta(2)$-position can thus not be excluded for the $[^{\bullet}AHA+H]^+$ ion.

The spectra recorded in H-stretching range generally show strong and sharp OH-stretch features and NH-stretch bands that are often weaker than predicted; this is likely due to the fact that OH moieties are free in the computed structures, whereas NH moieties, especially on the Lys side chain, engage in strong hydrogen bonds.[120-125] In any case, most pertinent to the present discussion, the spectra in this range provide little or no diagnostic information that would allow us to distinguish between isomers.

We finally inspect the cosine similarity (matching) scores between calculated and experimental spectra. The spectra predicted for the $C_\alpha(3)$ structures yield significantly higher values for all four systems, except for $[^{\bullet}AAK+H]^+$, for which $C_\alpha(3)$ and $C_\alpha(2)$ yield virtually equal scores. In conclusion, based on visual inspection of the spectra as well as on the cosine similarity value, the $C_\alpha(3)$ radical structures provide the best match, suggesting that radical migration does not take place after ETD for these systems under our experimental conditions. As the $C_\alpha(3)$ is not the lowest-energy structure, this also implies that the radical is kinetically trapped.

Figure 9-4 presents the experimental spectra for the two remaining z_3-ions studied in this work, $[^{\bullet}FAK+H]^+$ and $[^{\bullet}WAK+H]^+$. We compare again with calculated spectra for isomeric structures with the radical positioned at $C_\alpha(3)$, $C_\beta(3)$ or $C_\alpha(2)$. For these ions, radical migration from the cleavage site $C_\alpha(3)$ to the $C_\beta(3)$- or $C_\alpha(2)$-position yields similar energy gains of about 30 kJ/mol, so that based purely on thermochemical arguments no distinction can be made between these two. Earlier studies on ETD fragments containing Phe and Trp report calculations indicating that for $[^{\bullet}AFAR+H]^+$ z_4-ions, radical migration to the C_β position of Phe gives the lowest energy structure, while for the $[^{\bullet}AWAR+H]^+$ ion, the radical at the C_α of the Arg residue is the lowest-energy structure[80].

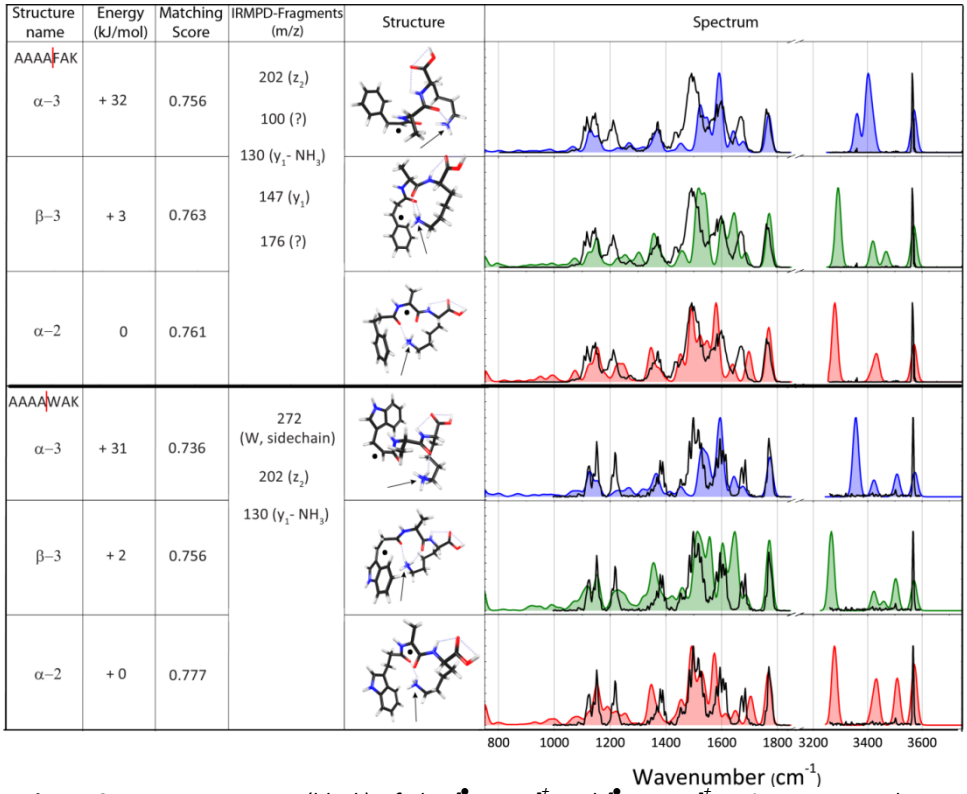


Figure 9-4. *IRMPD spectra (black) of the $[^{\bullet}FAK+H]^+$ and $[^{\bullet}WAK+H]^+$ z_3 ions compared to calculated spectra for isomers with the radical at the $C_\alpha(3)$ cleavage site (blue), at the $C_\beta(3)$ atom of the aromatic side chain (green) and at the $C_\alpha(2)$ atom enjoying captodative stabilization. The calculated spectra shown are the lowest-energy conformers after an MM/MD search for each of the isomers. In addition, structures with the radical position indicated, relative free energy in kJ/mol, IRMPD induced dissociation channels, and spectral matching scores based on the cosine similarity are given.*

Based on the cosine similarity scores of the experimental spectra with the calculated spectra, the $C_\alpha(2)$ isomer gives the most favorable match for both structures, but it is only marginally better than for the alternative structures. A visual inspection of experimental and computed spectra for the $C_\alpha(2)$ isomer shows a mismatch for both species at 1600 cm^{-1} (combined NH_3 bending, C=O stretching and N-H bending in the first amide bond). Also, the band around 1675 cm^{-1} is not reproduced in the calculated spectra for both species. However, calculated spectra for the alternative isomers do not reproduce this band satisfactorily either. A point in favor of the $C_\alpha(2)$ isomer is the observation of z_2-ions as IRMPD fragmentation channels, which may be an indication for radical migration to the $C_\alpha(2)$ position.

The calculated spectra for the $C_\beta(3)$-isomers in green show a doublet of peaks at about 1600 and 1650 cm^{-1}, similar to the $C_\beta(3)$-calculated spectrum for $[^\bullet YAK+H]^+$, which are not observed in the experimental spectra. This consistent mismatch appears to disqualify the $C_\beta(3)$-isomer for $[^\bullet FAK+H]^+$ and $[^\bullet WAK+H]^+$, although the IRMPD reaction channels for $[^\bullet WAK+H]^+$ include loss of the radical side chain, which may be an indication for the $C_\beta(3)$-isomer.

Despite the higher energy, the $C_\alpha(3)$-isomer with the radical at the cleavage site cannot be dismissed, certainly in light of our conclusions for the three systems in Figure 9-3. The main mismatch between computed and experimental spectra is observed for the intense experimental band around 1500 cm^{-1}, which is likely due to N-H bending in both amide bonds. This mismatch appears to be somewhat similar to that observed for the $[^\bullet AAK+H]^+$ ion in Figure 9-3, which otherwise gives a convincing match. Also, the band near 1200 cm^{-1} in the experimental spectra of $[^\bullet FAK+H]^+$ and $[^\bullet WAK+H]^+$ is not reproduced in the $C_\alpha(3)$ calculations, but all remaining peaks show a reasonably good agreement. We must conclude that a definitive assignment for these two ions remains insecure, but we tentatively assign a mixture of $C_\alpha(3)$ and $C_\alpha(2)$ isomers for the z_3-ions presented in Figure 9-4.

9.4. Conclusion

The position of the radical in z_3 ions generated by ETD on a set of six peptides has been investigated using a combination of infrared ion spectroscopy and quantum-chemical calculations. It is well known that (CID) MS/MS product ions may often be kinetically trapped species, so that an outright comparison of computed relative energies for different isomers has only limited value in predicting the structure of the product ions. Also in the present study, the z_3-ions with the radical positioned α-carbon of the third residue is a higher-energy isomer, with radical migration to the $C_\alpha(1)$, $C_\alpha(2)$ or $C_\beta(3)$ atoms leading to lower-energy isomers as a consequence of resonance stabilization, including captodative stabilization, of the radical. Nonetheless, we find that under the experimental conditions of our Bruker AmaZon ETD instrument, radical migration to form lower-energy isomers often does not occur for the z_3 ions investigated. Radical migration to the α-carbon of the C-terminal residue, $C_\alpha(1)$, can securely be ruled out based on the position of the carboxylic C=O stretch band. Further spectral comparison for the $[^\bullet AAK+H]^+$, $[^\bullet HAK+H]^+$, $[^\bullet YAK+H]^+$ and $[^\bullet AHA+H]^+$ z_3-ions, leads us to the conclusion that the radical remains kinetically trapped at the cleavage site; for $[^\bullet AHA+H]^+$ radical migration to the His $C_\beta(2)$ position cannot be excluded based on the IR spectrum. For the $[^\bullet FAK+H]^+$ and $[^\bullet WAK+H]^+$ z_3-ions, the current experimental and theoretical

spectra do not allow us to make a clear structural assignment; partial isomerization by radical migration to the adjacent amino acid residue, $C_\alpha(2)$, appears to occur. We tentatively attribute the experimental spectra to a mixture of isomers with the radical at the C_α-positions of the third and second residues.

Chapter 10

w-Type ions formed by electron transfer dissociation of cysteine containing peptides investigated by infrared ion spectroscopy

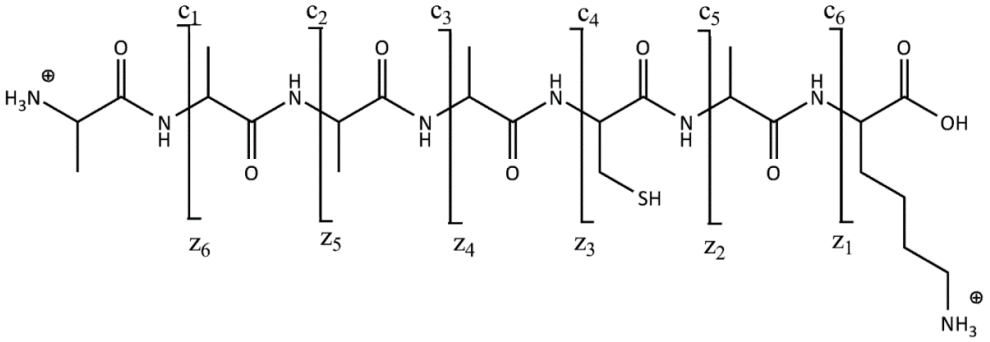
---

In mass spectrometry-based peptide sequencing, Electron Transfer Dissociation (ETD) and Electron Capture Dissociation (ECD) have become well established fragmentation methods complementary to collision-induced dissociation (CID). The dominant fragmentation pathways during ETD and ECD primarily involve the formation of c- and $z^{\bullet}$-type ions by cleavage of the peptide backbone at the N – C_{α} bond, although neutral losses from amino acid side chains have also been observed. Residue-specific neutral side chain losses provide useful information when conducting database searching and *de novo* sequencing. Here we use a combination of infrared ion spectroscopy and quantum-chemical calculations to assign the structures of two ETD-generated *w*-type fragment ions. These ions are spontaneously formed from ETD-generated $z^{\bullet}$-type fragments by neutral loss of 33 Da in peptides containing a cysteine residue. Analysis of the infrared ion spectra confirms that these $z^{\bullet}$-ions expel a thiol radical ($SH^{\bullet}$) and that a vinyl C=C group is formed at the cleavage site. $z^{\bullet}$-type fragments containing a Cys residue but not at the cleavage site do not spontaneously expel a thiol radical, but only upon additional collisional activation after ETD[*].

[*] Adapted from: Kempkes, L. J. M.; Martens, J.; Berden, G.; Oomens, J., w-Type ions formed by Electron Transfer Dissociation of Cys-containing peptides investigated by infrared ion spectroscopy. *J. Mass Spectrom.* ***2018****, 53,* 1207-1213.

10.1. Introduction

Mass spectrometry data in combination with bioinformatics algorithms form a key technology for peptide and protein sequencing, including the identification of post-translational protein modifications (PTMs).[126,127] Particularly with the advent of top-down proteomics and the sequencing of PTMs, electron induced dissociation methods – either electron capture or electron transfer dissociation (ECD or ETD) – have become popular.[10,11] The dominant ETD and ECD fragmentation pathways involve cleavage of the N-C_α bonds resulting in the formation of *c*- and $z^\bullet$-type ions [29,40,53,67] (see Scheme 10-1). Neutral loss fragments including those from the residue side chains are also commonly observed.[90,128-134] Side chain losses can confirm sequence assignments, improve database matching scores, and can be useful in *de novo* sequencing.[131] It is therefore useful to train sequencing algorithms to recognize neutral loss fragments in addition to sequence ions.[135] Neutral losses from residue side chains in radical ExD sequence ions have for instance been used to distinguish between leucine and isoleucine residues and to identify isomeric amino acid combinations.[135-138] On the other hand, extensive formation of non-sequence ions due to entire or partial side-chain losses, especially when remote from the backbone cleavage site, may severely complicate database searching.[139]



***Scheme 10-1**. Overview of expected ETD dissociation sites cleaving the peptide backbone at one of the N – C_α bonds, generating N-terminal c-ions and/or C-terminal z-ions. The z-ions are radical species.*

Mechanistically, neutral loss from radical $z^\bullet$-type ions has been explained by several charge-remote fragmentation (CRF) pathways.[129,139-141] Migration of the radical may occur from the original site of cleavage at the C_α atom to the C_β and C_γ positions of the same residue, or to C_α, C_β and C_γ atoms in neighboring residues.[34,131,136,138,139,142] Subsequent expulsion of a neutral radical generates even-electron fragment ions in the mass spectrum, which have been suggested to adopt a *w*-type fragment ion

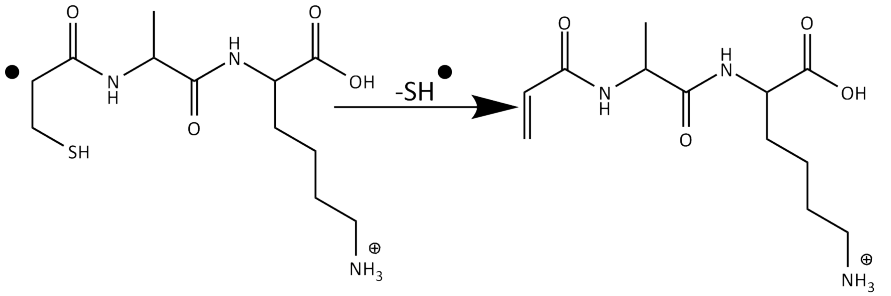
structure by cleavage of the C_β-C_γ bond of the N-terminal residue and the formation of a double bond between the C_α and C_β atoms (direct formation). Alternatively, cyclic *u*-type fragments may form as a result of bond formation between the N-terminal C_α atom and the C_β atom of the adjacent residue.[76,90,136,139] A possible mechanism for the loss of the entire side chain involves H-atom abstraction from the C_γ and subsequent α-cleavage and expulsion of the side chain as an even-electron neutral species, leaving the radical on the peptide backbone.[139]

Fung and Chan systematically investigated neutral loss product ions from ECD-generated $z^\bullet$-type ions using the doubly charged peptides $[RGGGXGGGR+2H]^{2+}$, where X denotes one of the 20 naturally occurring amino acid residues.[143] Residues that were shown to undergo secondary loss of a neutral radical, thus leaving behind an even-electron charged fragment, include among several others cysteine (Cys), the subject of investigation here. Even-electron neutral losses from residue side chains were also reported for various other amino acid residues.[67,84,142,143]

Cys-specific ExD pathways have been discussed in several studies.[82,141,143] For $z^\bullet$-type fragments from Cys-containing peptides, secondary fragmentation by loss of a 33 Da neutral fragment is common and is attributed to loss of a thiol radical ($SH^\bullet$). For z_6- and larger $z^\bullet$-type ions, neutral loss of CH_2S (46 Da) has also been observed.[143] ETD on $[AACAR+2H]^{2+}$ generated the z_4-ion (*m/z* 404) and the z_3-ion (*m/z* 333), where the intensity of the z_3-ion was significantly decreased by spontaneous loss of the SH radical, forming a secondary product at *m/z* 300.[82] ECD studies on peptides containing modified cysteine residues, such as carboxymethylated or carbamidomethylated Cys residues, report neutral losses corresponding to the modified Cys side chain.[144] The odd-electron neutral loss of $C_2H_4NOS^\bullet$ (90 Da) has also been observed in a negative electron-transfer dissociation (NETD) process on a doubly negatively charged ion containing carbamidomethylated Cys residues.[145]

In related studies not actually involving ExD, the CID fragmentation behavior, and possible radical migration in open-shell cations of Cys and Cys-containing peptides have been studied by theoretical[111,146,147] as well as experimental methods.[90,94,141,148-152] In experimental studies, CID-induced loss of nitric oxide (NO) from protonated S-nitrosocysteine is an efficient way to form radical cations of Cys and Cys-containing peptides.[148,149] Hao and Gross[149] showed that CID fragmentation of Cys-containing peptide radicals yields fragmentation similar to ExD experiments. Fragmentation reactions of radical cations of Cys-containing di- and tripeptides were found to be radical-driven or charge directed, leading to losses of $SH^\bullet$ and

CH_2S, respectively.[141] Analogous CID fragmentation products were found for $[GlyCysArg]^{\bullet+}$.[150] Observed CID products of the radical cation of the Cys amino acid were $HOCO^{\bullet}$, CH_2S, $CH_2SH^{\bullet}$, H_2S [148], while Cys-containing peptide radicals expel $SH^{\bullet}$ and CH_2S.[149] Theoretical investigations suggested that loss of a $HOCO^{\bullet}$ radical is energetically favored for the amino acid, followed by loss of the thiol radical and loss of the side chain ($CH_2SH^{\bullet}$).[111,146,147]



Scheme 10-2: *z-type ETD fragments expected for z-ions with a Cys residue at the cleavage site are not observed. Instead, a w-type ion lower in mass by 33 Da is observed due to additional neutral loss of a thiol radical. These w-type ions have generally been hypothesized to possess a vinyl group at the fragment's N-terminus.*

Although the vinyl structure of the *w*-type fragment ion has often been assumed (see Scheme 10-2), experimental verification of this structure is scarce. Here we use Infrared Multiple-Photon Dissociation (IRMPD) spectroscopy [58,59,61,63-65,75,99] to experimentally confirm their structure for two *z* – 33 ETD fragments from Cys containing peptides. Several previous reports have made use of IR and UV photodissociation action spectroscopy methods to characterize ETD and ECD product ion structures.[57,64,65,75,80,81,85,89,101-103,153]

10.2. Experimental and computational methods

10.2.1. IRMPD spectroscopy

The peptides AAAACAK and AAACAAK were purchased from BIOMATIK, Canada, and used without further purification. Doubly protonated peptide ions were generated by electrospray ionization (ESI) from 10^{-7} M solutions in 50:50 acetonitrile:water with 0.5% formic acid and mass-selectively stored in a modified 3D quadrupole ion trap mass spectrometer (Bruker, AmaZon Speed ETD).[58,75] The $z^{\bullet}$- and *w*-type fragment ions were produced via an ion-ion reaction with the fluoranthene radical anion for 300 ms. Ions of interest were then mass isolated in the trap and their IRMPD spectra were measured using the free electron laser FELIX

[116,154]. Ions were irradiated by two macropulses, each pulse having an energy of 10-60 mJ, at a repetition rate of 10 Hz and a bandwidth of ~0.5% of the center frequency. Relating parent and fragment ion intensities as the fragmentation yield (ΣI(fragment ions)/ΣI(parent + fragment ions)) and plotting the yield as a function of laser frequency generates an infrared spectrum. The yield is linearly corrected for frequency dependent variations in the laser power and the infrared frequency is calibrated using a grating spectrometer.

10.2.2. Computational chemistry

Optimized molecular geometries and theoretical infrared spectra were obtained using Density Functional Theory (DFT) calculations as implemented in Gaussian 09 revision D01[69]. The potential energy surface was explored to locate the lowest energy minima by using a molecular mechanics/molecular dynamics (MM/MD) approach employing AMBER 12[70]. Within AMBER an initial geometry optimization has been performed, followed by a simulated annealing procedure up to 500 K yielding 500 structures. These structures were reduced to about 20 to 30 candidate structures by considering their structural similarity setting appropriate rms criteria. In the final step, these candidate structures were optimized using DFT and their harmonic vibrational spectra were calculated. All computed harmonic vibrational frequencies were scaled by 0.975 and convoluted with a 25 cm^{-1} full-width-at-half-maximum (FWHM) Gaussian line shape to facilitate comparison with experimental spectra. Structure optimization, thermodynamic corrections and frequency calculations were performed using B3LYP/6-31++G(d,p). Geometries and relative energies of conformers were verified at the M06-2X/6-31++G(d,p) level of theory. Single-point electronic energies were also obtained at the MP2(full)/6-31+G(d,p)//B3LYP/6-31++G(d,p) level for comparison. The computational procedure is described in more detail elsewhere.[59,71,155]

10.3. Results and discussion

Figure 10-1 shows the ETD mass spectra for the two doubly protonated heptapeptides AAAACAK and AAACAAK investigated in this study, having a Cys residue in the third and fourth position from the C-terminus, respectively. The identified ETD fragment ions are annotated in Figure 10-1. As a consequence of the basic Lys residue at the C-terminus, the series of $z^{\bullet}$-type fragment ions are nearly complete, except for the smallest ions

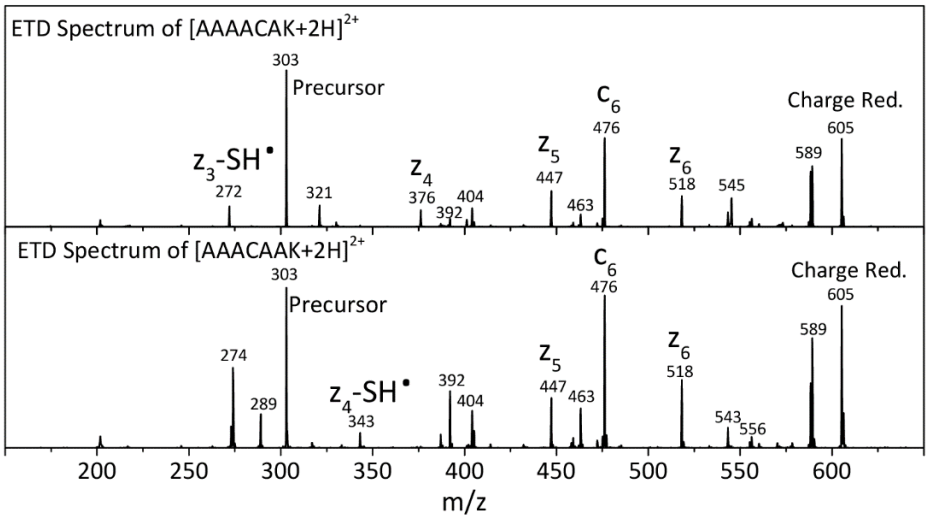


Figure 10-1. *ETD mass spectra of $[AAAACAK+2H]^{2+}$ (top panel) and $[AAACAAK+2H]^{2+}$ (bottom panel). ETD fragment ions are annotated. Neutral loss of an $SH^{\bullet}$ radical is observed for the $z_3^{\bullet}$ fragment from AAAACAK and for the $z_4^{\bullet}$ ion from AAACAAK, i.e. from $z^{\bullet}$-ions with the Cys residue at the cleavage site.*

Interestingly, for AAAACAK – with Cys in the third position from the C-terminus – the z_3 fragment is missing, but instead a peak 33 Da lower in mass is observed. Analogously, for AAACAAK the z_4 ion is missing and instead a peak at the nominal mass of z_4 – 33 is observed. We observe that neutral loss of a thiol radical ($SH^{\bullet}$) occurs only for $z^{\bullet}$-ions that contain a cysteine residue at the cleavage site; these fragments are commonly referred to as *w*-type ions. For the longer $z^{\bullet}$-ions, which do contain a Cys residue but not at the cleavage site, $SH^{\bullet}$-loss is not observed upon ETD. However, isolating one of these $z^{\bullet}$-ions and applying CID in an MS^3 experiment did result in expulsion of neutral fragments of 33 Da and 46 Da, the latter likely corresponding to loss of CH_2S (see Figures 10-2 and 10-3).

The $z^{\bullet}$-type ETD fragment ions that underwent neutral loss of 33 Da were investigated using IRMPD spectroscopy to determine their molecular structures. Upon IRMPD of the z_3 – 33 ETD fragment ion, *i.e.* the w_3-ion, dissociation is observed into *m/z* 147 (y_1-ion), *m/z* 129 (Lys-residue) and *m/z* 84 (fragment of Lys residue). For the w_4 ETD fragment ion, IRMPD leads to fragmentation channels *m/z* 218 (y_2-ion), *m/z* 147 (y_1-ion) and *m/z* 129 (Lys-residue). IRMPD spectra are constructed including fragmentation into all of these mass channels.

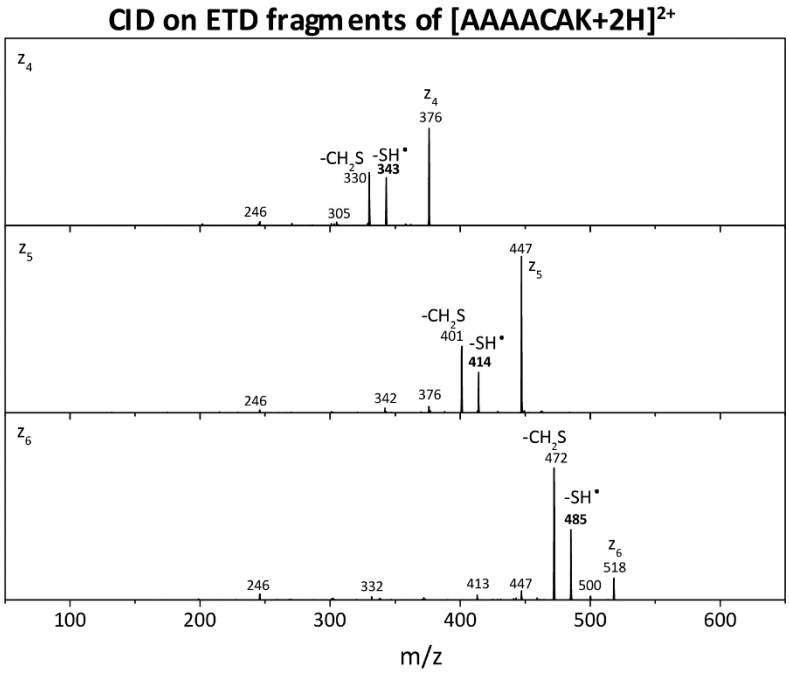


Figure 10-2. *Mass spectra of CID applied on z•-ions obtained from ETD of [AAAACAK+2H]2+. The mass peaks corresponding to loss of the radical thiol group are indicated.*

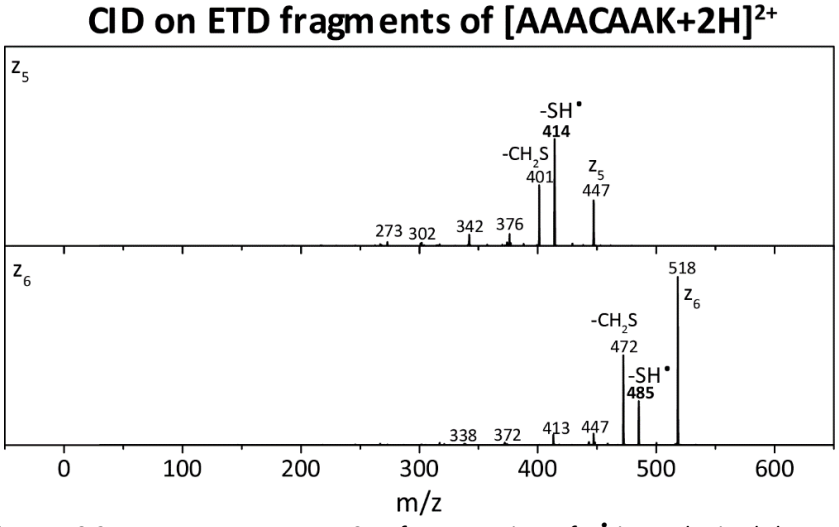


Figure 10-3. *Mass spectra upon CID fragmentation of z•-ions obtained by ETD on [AAACAAK+2H]2+. The mass peaks corresponding to loss of the radical thiol group are indicated.*

The left panel of Figure 10-4 shows the IRMPD spectrum of the w_3 fragment ion from $[AAAACAK+2H]^{2+}$ at *m/z* 272, and that of the w_4 fragment ion from $[AAACAAK+2H]^{2+}$ at *m/z* 343 is shown in the right panel. For both these z_3 – 33 and the z_4 – 33 ions, computed spectra for a structure containing a vinyl group at the N-terminus match favorably with the experimental spectra. In this structure, the radical thiol group is expelled from the cysteine side chain and a double bond is formed between the C_α and C_β atoms at the site of cleavage. The calculated spectra of the lowest energy conformers (red) and the best matching conformers (blue), from a visual inspection, are presented. Computed spectra for additional conformers of these species are shown in the Supporting Information Figure S3 of ref. [115].

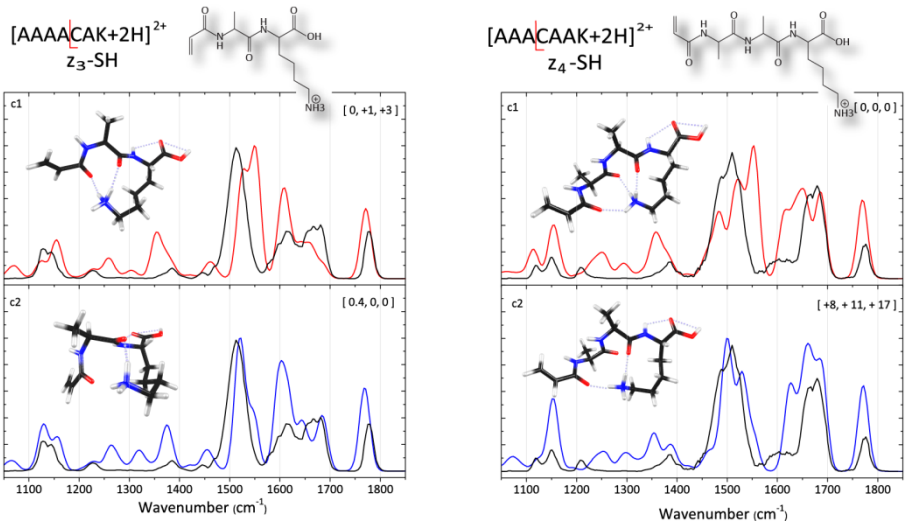


Figure 10-4. *IRMPD spectra of the z_3 – 33 ETD fragment (w_3) from $[AAAACAK+2H]^{2+}$ at m/z 272 (black, left panels) and of the z_4 – 33 ETD fragment (w_4) of $[AAACAAK+2H]^{2+}$ at m/z 343 (black, right panels). The experimental spectra are compared to predicted spectra for the lowest energy structures (red) and to the qualitatively best matching calculated spectra (blue). Relative energies (kJ/mol) at the B3LYP, M06-2X and MP2(full) levels of theory are given in the square brackets.*

For the w_3-ion of $[AAAACAK+2H]^{2+}$, the calculated spectra of two effectively iso-energetic conformers c1 and c2 are shown in the two panels on the left in Figure 10-4. Both conformers are compact structures where the protonated Lys amine group hydrogen-bonds with the two amide carbonyl oxygen atoms of the backbone. The band near 1770 cm^{-1} is attributed to the carbonyl stretch of the C-terminal carboxylic acid group. In the computed spectra of both conformers, the

unresolved feature between 1600 and 1700 cm^{-1} consists of four relatively strong bands having amide C=O stretch character mixed with ammonium bending character. In addition, the band near 1660 cm^{-1} in both conformers has significant vinyl C=C stretch character mixed in. The intense feature just above 1500 cm^{-1} is due to three vibrational modes, symmetric ammonium bending and the two amide NH bending modes. The c1 calculation attributes the highest intensity to the ammonium bending mode near 1550 cm^{-1}, while the c2 calculation attributes the highest intensity to the amide NH bending modes around 1520 cm^{-1}, leading to a better match for c2. The band at 1355 cm^{-1} in c1 (mainly delocalized CH bending character) is shifted to 1375 cm^{-1} in conformer c2, the latter being in slightly better agreement with experiment.

To the right in Figure 10-4, two different conformers for the vinyl group containing structure of the w_4-ion from $[AAACAAK+2H]^{2}$ are shown along with their predicted IR spectra. The lowest-energy conformer (c1), in the top panel has its protonated Lys amine group hydrogen-bonded with all three amide oxygen atoms. The bottom panel displays a higher-energy conformer (c2) in which the ammonium group is hydrogen-bonded with just two amide oxygen atoms; the central amide moiety in the molecule is not engaged in hydrogen bonding. This leads to differences in the two predicted spectra around 1650-1600 cm^{-1} and 1550-1480 cm^{-1}. Although higher in energy at all levels of theory considered here, the c2 conformer qualitatively provides a better spectral match and we suggest to assign this structure.

The band near 1770 cm^{-1} is again the C-terminal COOH carbonyl stretch. The feature roughly observed between 1600 and 1700 cm^{-1}, including the shoulder on the low-frequency side, is due to several normal modes mixing amide CO stretch character with asymmetric bending of the Lys NH_3^+ group. Again, the vinyl C=C stretch is mixed with the CO stretch of the adjacent amide and this mode appears at 1622 cm^{-1} in c2 and 1635 cm^{-1} in c1, at the position of the shoulder in the experimental spectrum. The c1 calculation predicts a strong, quite red-shifted CO stretch/ NH_3^+ bending mode at 1613 cm^{-1}, which causes the c1 conformer to match poorly with experiment. Also the amide II experimental feature around 1500 cm^{-1} matches better with the computed spectrum of c2 than with that of c1. In both spectra, the feature comprises four intense bands, the symmetric NH_3^+ bending mode and three amide NH bending modes. The ammonium bending mode is predicted near 1550 cm^{-1} for both conformers, but is four times more intense in c1 than in c2, explaining the better match with experiment for c2. The remainder of the spectrum is analogous to that of the w_3-ion, with the C-terminal COH bending mode near 1150 cm^{-1} as the most prominent peak.

Secondary neutral loss from $z^{\bullet}$-type ions from ExD – and thiol loss from $z^{\bullet}$-type ions containing Cys residues in particular – leading to *w*-type ions has been addressed in various investigations. [90,131,136,139,144] The structure of *w*-type ions are commonly assumed to involve a double bond between the α- and β-carbon atoms of the residue at the cleavage site. This fragment ion structure was hypothesized on the basis of differences in the CID MS/MS behavior of $z^{\bullet}$-ions versus (z+1)-ions.[138] Related, the differences in bond dissociation energies for H-atom removal from the α and β carbons have been used to rationalize the CC double bond structure of the *w*-ion.[90,139] Computational investigations of the potential energy surface of the dissociation reaction at the DFT level also indicate the formation of a vinyl group at the N-terminus of the *w*-ion[143]. However, direct experimental probing of the *w*-ion structure has until now not been reported to our knowledge. In this contribution, IR spectroscopy of two *w*-type ions formed by loss of an SH-radical from a Cys residue at the cleavage site confirms the vinyl group structure suggested previously. [90,131,136,139,143,144] Indirectly, the observation of this structure also confirms the suggested charge remote fragmentation pathways suggested. [139-141,143,144]

Other $z^{\bullet}$-type ions were observed to expel a thiol radical only upon applying additional CID activation after ETD, in line with reported higher rates of formation of *w*-ions compared to *u*-ions.[136] Also in accordance with previous studies,[82,143] loss of CH_2S (46 Da) after collisional activation was only observed for $z^{\bullet}$-ions which contain a Cys residue, but not at the cleavage site. Spectroscopic probing of the structure of these fragment ions is challenging because of the more involved ETD/CID MS^n scheme required to produce them and the inherent lower ion yields as compared to the *w*-type ions; such experiments are reserved for future studies.

10.4. Conclusion

We have reported the first infrared spectra for *w*-type fragment ions resulting from ETD of two Cys-containing peptides. Specifically, the mass peak at *m/z* 272 in the ETD mass spectrum of $[AAAACAK+2H]^{2+}$ and that at *m/z* 343 in the ETD mass spectrum of $[AAACAAK+2H]^{2+}$, which cannot be assigned to sequence ions but are the result of neutral $SH^{\bullet}$-loss from the radical z_3 and z_4-ions of the respective peptides, have been investigated with infrared ion spectroscopy. Evaluation of experimental and predicted IR spectra confirm that these *w*-type fragment ions possess a structure having a C=C double bond at their N-terminus. This structure supports the mechanism proposed for the formation of *w*-type fragment ions in ExD, proceeding through cleavage of the C_β-S bond of the Cys residue with

concomitant formation of a double bond between the C_α and C_β atoms of the *w*-ion. Elimination of the $SH^\bullet$-group is only observed for $z^\bullet$-ions with a cysteine residue at the cleavage site, independent of its position of cysteine in the peptide.

Bibliography

[1] Hunt, D.F., Yates, J.R., Shabanowitz, J., Winston, S., Hauer, C.R., *Proc. Natl. Acad. Sci. USA*, **83**, (1986), 6233-6237.
[2] Wells, J.M., McLuckey, S.A., *Methods Enzymol.*, **402**, (2005), 148-185.
[3] Paizs, B., Suhai, S., *Mass Spectrom. Rev.*, **24**, (2005), 508-548.
[4] Cooper, H.J., Håkansson, K., Marshall, A.G., *Mass Spectrom. Rev.*, **24**, (2005), 201-222.
[5] Wiesner, J., Premsler, T., Sickmann, A., *Proteomics*, **8**, (2008), 4466-4483.
[6] Mirgorodskaya, E., Roepstorff, P., Zubarev, R.A., *Anal. Chem.*, **71**, (1999), 4431-4436.
[7] Stensballe, A., Jensen, O.N., Olsen, J.V., Haselmann, K.F., Zubarev, R.A., *Rapid Commun. Mass Spectrom.*, **14**, (2000), 1793-1800.
[8] Håkansson, K., Cooper, H.J., Emmett, M.R., Costello, C.E., Marshall, A.G., Nilsson, C.L., *Anal. Chem.*, **73**, (2001), 4530-4536.
[9] Mikesh, L.M., Ueberheide, B., Chi, A., Coon, J.J., Syka, J.E.P., Shabanowitz, J., Hunt, D.F., *Biochim. Biophys. Acta Proteins Proteomics*, **1764**, (2006), 1811-1822.
[10] Riley, N.M., Coon, J.J., *Anal. Chem.*, **90**, (2018), 40-64.
[11] Coon, J.J., *Anal. Chem.*, **81**, (2009), 3208-3215.
[12] Swaney, D.L., McAlister, G.C., Coon, J.J., *Nat. Meth.*, **5**, (2008), 959-964.
[13] Breuker, K., McLafferty, F.W., *Angew. Chem. Int. Ed.*, **42**, (2003), 4900-4904.
[14] Breuker, K., Oh, H., Lin, C., Carpenter, B.K., McLafferty, F.W., *Proc. Natl. Acad. Sci. USA*, **101**, (2004), 14011-14016.
[15] Oh, H., Breuker, K., Sze, S.K., Ge, Y., Carpenter, B.K., McLafferty, F.W., *Proc. Natl. Acad. Sci. USA*, **99**, (2002), 15863-15868.
[16] Breuker, K., McLafferty, F.W., *Proc. Natl. Acad. Sci. USA*, **105**, (2008), 18145-18152.
[17] Zubarev, R.A., Horn, D.M., Fridriksson, E.K., Kelleher, N.L., Kruger, N.A., Lewis, M.A., Carpenter, B.K., McLafferty, F.W., *Anal. Chem.*, **72**, (2000), 563-573.
[18] Coon, J.J., Ueberheide, B., Syka, J.E.P., Dryhurst, D.D., Ausio, J., Shabanowitz, J., Hunt, D.F., *Proc. Natl. Acad. Sci. USA*, **102**, (2005), 9463-9468.
[19] Kaplan, D.A., Hartmer, R., Speir, J.P., Stoermer, C., Gumerov, D., Easterling, M.L., Brekenfeld, A., Kim, T., Laukien, F., Park, M.A., *Rapid Commun. Mass Spectrom.*, **22**, (2008), 271-278.
[20] Kelleher, N.L., Zubarev, R.A., Bush, K., Furie, B., Furie, B.C., McLafferty, F.W., Walsh, C.T., *Anal. Chem.*, **71**, (1999), 4250-4253.
[21] Chi, A., Huttenhower, C., Geer, L.Y., Coon, J.J., Syka, J.E.P., Bai, D.L., Shabanowitz, J., Burke, D.J., Troyanskaya, O.G., Hunt, D.F., *Proc. Natl. Acad. Sci. USA*, **104**, (2007), 2193-2198.
[22] Molina, H., Horn, D.M., Tang, N., Mathivanan, S., Pandey, A., *Proc. Natl. Acad. Sci. USA*, **104**, (2007), 2199-2204.

[23] Stephenson Jr, J.L., McLuckey, S.A., Reid, G.E., Wells, J.M., Bundy, J.L., *Curr. Opinion Biotechnol.*, **13**, (2002), 57-64.
[24] Syka, J.E.P., Coon, J.J., Schroeder, M.J., Shabanowitz, J., Hunt, D.F., *Proc. Nat. Acad. Sci. USA*, **101**, (2004), 9528-9533.
[25] Jones, J.W., Sasaki, T., Goodlett, D.R., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **18**, (2007), 432-444.
[26] Pepin, R., Tureček, F., *J. Phys. Chem. B.*, **119**, (2015), 2818-2826.
[27] Tureček, F., *J. Am. Chem. Soc.*, **125**, (2003), 5954-5963.
[28] Tsybin, Y.O., He, H., Emmett, M.R., Hendrickson, C.L., Marshall, A.G., *Anal. Chem.*, **79**, (2007), 7596-7602.
[29] Zubarev, R.A., Kelleher, N.L., McLafferty, F.W., *J. Am. Chem. Soc.*, **120**, (1998), 3265-3266.
[30] Kruger, N.A., Zubarev, R.A., Horn, D.M., McLafferty, F.W., *Int. J. Mass Spectrom.*, **185-187**, (1999), 787-793.
[31] Zubarev, R.A., *Mass Spectrom. Rev.*, **22**, (2003), 57-77.
[32] Zubarev, R.A., Zubarev, A.R., Savitski, M.M., *J. Am. Soc. Mass Spectrom.*, **19**, (2008), 753-761.
[33] Savitski, M.M., Kjelden, F., Nielsen, M.L., Zubarev, R.A., *Angew. Chem. Int. Ed.*, **45**, (2006), 5301-5303.
[34] Tureček, F., Julian, R.R., *Chem. Rev.*, **113**, (2013), 6691-6733.
[35] Zhurov, K.O., Fornelli, L., Wodrich, M.D., Laskay, U.A., Tsybin, Y.O., *Chem. Soc. Rev.*, **42**, (2013), 5014-5030.
[36] McLuckey, S.A., Stephenson Jr, J.L., *Mass Spectrom. Rev.*, **17**, (1998), 369-407.
[37] Martens, J., Berden, G., Oomens, J., *Anal. Chem.*, **88**, (2016), 6126-6129.
[38] Tureček, F., Syrstad, E.A., *J. Am. Chem. Soc.*, **125**, (2003), 3353-3369.
[39] Zubarev, R.A., Kruger, N.A., Fridriksson, E.K., Lewis, M.A., Horn, D.M., Carpenter, B.K., McLafferty, F.W., *J. Am. Chem. Soc.*, **121**, (1999), 2857-2862.
[40] Syrstad, E.A., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **16**, (2005), 208-224.
[41] Sawicka, A., Skurski, P., Hudgins, R.R., Simons, J., *J. Phys. Chem. B*, **107**, (2003), 13505-13511.
[42] Syrstad, E.A., Stephens, D.D., Tureček, F., *J. Phys. Chem. A.*, **107**, (2003), 115-126.
[43] Syrstad, E.A., Tureček, F., *J. Phys. Chem. A.*, **105**, (2001), 11144-11155.
[44] Anusiewicz, I., Berdys-Kochanska, J., Simons, J., *J. Phys. Chem. A.*, **109**, (2005), 5801-5813.
[45] Anusiewicz, I., Berdys-Kochanska, J., Skurski, P., Simons, J., *J. Phys. Chem. A.*, **110**, (2006), 1261-1266.
[46] Sobczyk, M., Simons, J., *J. Phys. Chem. B.*, **110**, (2006), 7519-7527.
[47] Chen, X., Tureček, F., *J. Am. Chem. Soc.*, **128**, (2006), 12520-12530.
[48] Simons, J., *Chem. Phys. Lett.*, **484**, (2010), 81-95.
[49] Patriksson, A., Adams, C., Kjeldsen, F., Raber, J., van der Spoel, D., Zubarev, R.A., *Int. J. Mass Spectrom.*, **248**, (2006), 124-135.

[50] Wodrich, M.D., Zhurov, K.O., Vorobyev, A., Hamidane, H.B., Corminboeuf, C., Tsybin, Y.O., *J. Phys. Chem. B.*, **116**, (2012), 10807-10815.
[51] Wodrich, M.D., Zhurov, K.O., Corminboeuf, C., Tsybin, Y.O., *J. Phys. Chem. B.*, **118**, (2014), 2985-2992.
[52] Schneeberger, E.-M., Breuker, K., *Chemical Science*, **9**, (2018), 7338-7353.
[53] Zubarev, R.A., Haselmann, K.F., Budnik, B., Kjeldsen, F., Jensen, F., *Eur. J. Mass Spectrom.*, **8**, (2002), 337-349.
[54] Sobczyk, M., Anusiewicz, I., Berdys-Kochanska, J., Sawicka, A., Skurski, P., Simons, J., *J. Phys. Chem. A.*, **109**, (2005), 250-258.
[55] Tureček, F., Transient Intermediates of Chemical Reactions by Neutralization-Reionization Mass Spectrometry, in: C.A. Schalley (Ed.) Modern Mass Spectrometry, Springer Berlin Heidelberg, Berlin, Heidelberg, 2003, pp. 77-129.
[56] Anusiewicz, I., Skurski, P., Simons, J., *J. Phys. Chem. B.*, **118**, (2014), 7892-7901.
[57] Imaoka, N., Houferak, C., Murphy, M.P., Nguyen, H.T.H., Dang, A., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **29**, (2018), 1768-1780.
[58] Martens, J., Berden, G., Gebhardt, C.R., Oomens, J., *Rev. Sci. Instrum.*, **87**, (2016), 103108.
[59] Kempkes, L.J.M., Martens, J., Grzetic, J., Berden, G., Oomens, J., *J. Am. Soc. Mass Spectrom.*, **27**, (2016), 1855-1869.
[60] Lucas, B., Grégoire, G., Lemaire, J., Maître, P., Ortega, J., Rupenyan, A., Reimann, B., Schermann, J.P., Desfrançois, C., *Phys. Chem. Chem. Phys.*, **6**, (2004), 2659-2663.
[61] Polfer, N.C., Oomens, J., *Phys. Chem. Chem. Phys.*, **9**, (2007), 3804-3817.
[62] Fung, Y.M.E., Besson, T., Lemaire, J., P.;, M., Zubarev, R.A., *Angew. Chem. Int. Ed. Engl.*, **48**, (2009), 8340-8342.
[63] Correia, C.F., Balaj, P.O., Scuderi, D., Maitre, P., Ohanessian, G., *J. Am. Chem. Soc,* **130**, (2008), 3359-3370.
[64] Frison, G., van der Rest, G., Tureček, F., Besson, T., Lemaire, J., Maître, P., Chamot-Rooke, J., *J. Am. Chem. Soc*, **130**, (2008), 14916-14917.
[65] Shaffer, C.J., Martens, J., Marek, A., Oomens, J., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **27**, (2016), 1176-1185.
[66] Moss, C.L., Chamot-Rooke, J., Nicol, E., Brown, J., Campuzano, I., Richardson, K., Williams, J.P., Bush, M.F., Bythell, B., Paizs, B., Tureček, F., *J. Phys. Chem. B.*, **116**, (2012), 3445-3456.
[67] Han, H., Xia, Y., McLuckey, S.A., *J. Proteome Res.*, **6**, (2007), 3062-3069.
[68] Pepin, R., Laszlo, K.J., Peng, B., Marek, A., Bush, M.F., Tureček, F., *J. Phys. Chem. A.*, **118**, (2014), 308-324.
[69] Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G.A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H.P., Izmaylov, A.F., Bloino, J., Zheng, G., Sonnenberg, J.L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J., J.A.;, Peralta, J.E., Ogliaro, F., Bearpark, M., Heyd, J.J., Brothers, E., Kudin, K.N.,

Staroverov, V.N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Rega, N., Millam, N.J., Klene, M., Knox, J.E., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Martin, R.L., Morokuma, K., Zakrzewski, V.G., Voth, G.A., Salvador, P., Dannenberg, J.J., Dapprich, S., Daniels, A.D., Farkas, O., Foresman, J.B., Ortiz, J.V., Cioslowski, J., Fox, D.J., in: Gaussian, Inc., Wallingford, CT, Gaussian, Inc., Wallingford, CT, 2009.
[70] Case, D.A., Darden, T., Cheatham III, T.E., Simmerling, C., Wang, J., Duke, R.E., Luo, R., Walker, R.C., Zhang, W., Merz, K.M., Roberts, B.P., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Goetz, A.W., Kolossváry, I., Wong, K.F., Paesani, F., Vanicek, J., Wolf, R.M., Liu, J., Wu, X., Brozell, S.R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D.R., Mathews, D.H., Seetin, M.G., Salomon-Ferrer, R., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., Kollman, P.A., (2012).
[71] Martens, J., Grzetic, J., Berden, G., Oomens, J., *Int. J. Mass Spectrom.* , **377**, (2015), 179-187.
[72] Grzetic, J., Oomens, J., *Int. J. Mass Spectrom.*, **316-318**, (2012), 216-226.
[73] Bythell, B.J., *J. Phys. Chem. A.*, **118**, (2014), 10797-10803.
[74] Schennach, M., Breuker, K., *J. Am. Soc. Mass Spectrom.*, **26**, (2015), 1059-1067.
[75] Martens, J., Grzetic, J., Berden, G., Oomens, J., *Nature Commun.*, **7**, (2016), 11754.
[76] Kjeldsen, F., Haselmann, K.F., Budnik, B.A., Jensen, F., Zubarev, R.A., *Chem. Phys. Lett.*, **356**, (2002), 201-206.
[77] Chung, T.W., Hui, R., Ledvina, A.R., Coon, J.J., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **23**, (2012), 1336-1350.
[78] Ledvina, A.R., Coon, J.J., Tureček, F., *Int. J. Mass Spectrom.*, **377**, (2015), 44-53.
[79] Pepin, R., Layton, E.D., Liu, Y., Afonso, C., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **28**, (2017), 164-181.
[80] Nguyen, H.T.H., Shaffer, C.J., Pepin, R., Tureček, F., *J. Phys. Chem. Lett.*, **6**, (2015), 4722-4727.
[81] Nguyen, H.T.H., Shaffer, C.J., Tureček, F., *J. Phys. Chem. B.*, **119**, (2015), 3948-3961.
[82] Nguyen, H.T.H., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **28**, (2017), 1333-1344.
[83] Riffet, V., Jacquemin, D., Frison, G., *Int. J. Mass Spectrom.*, **390**, (2015), 28-38.
[84] Chung, T.W., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **21**, (2010), 1279-1295.
[85] Viglino, E., Shaffer, C.J., Tureček, F., *Angew. Chem. Int. Ed.*, **55**, (2016), 7469-7473.
[86] Piatkivskyi, A., Osburn, S., Jaderberg, K., Grzetic, J., Steill, J.D., Oomens, J., Zhao, J., Lau, J.K.-C., Verkerk, U.H., Hopkinson, A.C., Siu, K.W.M., Ryzhov, V., *J. Am. Soc. Mass Spectrom.*, **24**, (2013), 513-523.
[87] Bellina, B., Compagnon, I., Houver, S., Maitre, P., Allouche, A.-R., Antoine, R., Dugourd, P., *Angew. Chem. Int. Ed.*, **50**, (2011), 11430-11430.

[88] O'Connor, P.B., Lin, C., Cournoyer, J.J., Pittman, J.L., Belyayev, M., Budnik, B.A., *J. Am. Soc. Mass Spectrom.*, **17**, (2006), 576-585.
[89] Nguyen, H.T.H., Andrikopoulos, P.C., Bim, D., Rulisek, L., Dang, A., Tureček, F., *J. Phys. Chem. B.*, **121**, (2017), 6557-6569.
[90] Sun, Q., Nelson, H., Ly, T., Stoltz, B.M., Julian, R.R., *J. Proteome Res.*, **8**, (2009), 958-966.
[91] Ly, T., Julian, R.R., *J. Am. Soc. Mass Spectrom.*, **20**, (2009), 1148-1158.
[92] Moore, B.N., Blanksby, S.J., Julian, R.R., *Chem. Commun.*, (2009), 5015-5017.
[93] Zhang, X., Julian, R.R., *J. Am. Soc. Mass Spectrom.*, **24**, (2013), 524-533.
[94] Osburn, S., Berden, G., Oomens, J., O'Hair, R.A.J., Ryzhov, V., *J. Am. Soc. Mass Spectrom.*, **22**, (2011), 1794-1803.
[95] Zhao, J., Song, T., Xu, M., Quan, Q., Siu, K.W.M., Hopkinson, A.C., Chu, I.K., *Phys. Chem. Chem. Phys.*, **14**, (2012), 8723-8791.
[96] Mu, X., Song, T., Xu, M., Lai, *J. Phys. Chem. B*, **118**, (2014), 4273-4281.
[97] Chu, I.K., Zhao, J., Xu, M., Siu, S.O., Hopkinson, A.C., Siu, K.W.M., *J. Am. Chem. Soc*, **130**, (2008), 7862-7872.
[98] Steill, J.D., Zhao, J., Siu, C., Ke, Y., Verkerk, U.H., Oomens, J., Dunbar, R.C., Hopkinson, A.C., Siu, K.W.M., *Angew. Chem. Int. Ed.*, **47**, (2008), 9666-9668.
[99] Ledvina, A.R., Chung, T.W., Hui, R., Coon, J.J., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **23**, (2012), 1351-1363.
[100] Moore, B.N., Ly, T., Julian, R.R., *J. Am. Chem. Soc.*, **133**, (2011), 6997-7006.
[101] Viglino, E., Lai, C.K., Mu, X., Chu, I.K., Tureček, F., *J. Am. Soc. Mass Spectrom.*, **27**, (2016), 1454-1467.
[102] Shaffer, C.J., Pepin, R., Tureček, F., *J. Mass Spectrom.*, **50**, (2015), 1438-1442.
[103] Nguyen, H.T.H., Shaffer, C.J., Ledvina, A.R., Coon, J.J., Tureček, F., *Int. J. Mass Spectrom.*, **378**, (2015), 20-30.
[104] Swaney, D.L., McAlister, G.C., Wirtala, M., Schwartz, J.C., Syka, J.E.P., Coon, J.J., *Anal. Chem.*, **79**, (2007), 477-485.
[105] Xia, Y., Han, H., McLuckey, S.A., *Anal. Chem.*, **80**, (2008), 1111-1117.
[106] Liu, J., Liang, X., McLuckey, S.A., *J. Proteome Res.*, **7**, (2008), 130-137.
[107] Mu, X., Song, T., Siu, C.-K., Chu, I.K., *The Chemical Record*, **18**, (2018), 20-44.
[108] Kempkes, L.J.M., Martens, J., Berden, G., Oomens, J., *J. Phys. Chem. Lett.*, **9**, (2018), 6404-6411.
[109] Viehe, H.G., Janousek, Z., Merenyi, R., Stella, L., *Acc. Chem. Res.*, **18**, (1985), 148-154.
[110] Bagheri-Majdi, E., Ke, Y., Orlova, G., Chu, I.K., Hopkinson, A.C., Siu, K.W.M., *J. Phys. Chem. B.*, **108**, (2004), 11170-11181.
[111] Zhao, J., Siu, K.W.M., Hopkinson, A.C., *Phys. Chem. Chem. Phys.*, **10**, (2008), 281-288.
[112] Croft, A.K., Easton, C.J., Radom, L., *J. Am. Chem. Soc*, **125**, (2003), 4119-4124.
[113] Burgers, P.C., Holmes, J.L., Terlouw, J.K., van Baar, B., *Org. Mass Spectrom.*, **20**, (1985), 202-206.

[114] Siu, C., Ke, Y., Orlova, G., Hopkinson, A.C., Siu, K.W.M., *J. Am. Soc. Mass Spectrom.*, **19**, (2008), 1799-1807.
[115] Kempkes, L.J.M., Martens, J., Berden, G., Oomens, J., *J. Mass. Spectrom.*, **53**, (2018), 1207-1213.
[116] Oepts, D., van der Meer, A.F.G., van Amersfoort, P.W., *Infrared Phys. Technol.*, **36**, (1995), 297-308.
[117] Kempkes, L.J.M., Martens, J., Berden, G., Oomens, J., *Int. J. Mass Spectrom.*, **429**, (2018), 90-100.
[118] Fu, W., Hopkins, W.S., *J. Phys. Chem. A* **122**, (2018), 167-171.
[119] Kempkes, L.J.M., Martens, J., Berden, G., Houthuijs, K.J., Oomens, J., *Faraday Discuss.*, (2018).
[120] Avilés-Moreno, J.R., Berden, G., Oomens, J., Martinez-Haya, B., *Phys. Chem. Chem. Phys.*, **20**, 8968-8975.
[121] Gao, J., Berden, G., Rodgers, M.T., Oomens, J., *Phys. Chem. Chem. Phys.*, **18**, (2016), 7269-7277.
[122] Yang, B., Wu, R.R., Berden, G., Oomens, J., Rodgers, M.T., *J. Phys. Chem. B*, **117**, (2013), 14191-14201.
[123] Scuderi, D., Bakker, J.M., Durand, S., Maitre, P., Sharma, A., Martens, J.K., Nicol, E., Clavaguéra, C., Ohanessian, G., *Int. J. Mass Spectrom.*, **308**, (2011), 338-347.
[124] DeBlase, A.F., Bloom, S., Lectka, T., Jordan, K.D., McCoy, A.B., Johnson, M.A., *J. Chem. Phys.*, **139**, (2013), 024301.
[125] Scutelnic, V., Perez, M.A.S., Marianski, M., Warnke, S., Gregor, A., Rothlisberger, U., Bowers, M.T., Baldauf, C., von Helden, G., Rizzo, T.R., Seo, J., *J. Am. Chem. Soc.*, **140**, (2018), 7554-7560.
[126] Srikanth, R., Wilson, J., Bridgewater, J.D., Numbers, J.R., Lim, J., Olbris, M.R., Kettani, A., Vachet, R.W., *J. Am. Soc. Mass Spectrom.*, **18**, (2007), 1499-1506.
[127] Kandasamy, K., Pandey, A., Molina, H., *Anal. Chem.*, **81**, (2009), 7170-7180.
[128] Cooper, H.J., Hudgins, R.R., Håkansson, K., Marshall, A.G., *J. Am. Soc. Mass Spectrom.*, **13**, (2002), 241-249.
[129] Leymarie, N., Costello, C.E., O'Connor, P.B., *J. Am. Chem. Soc.*, **125**, (2003), 8949-8958.
[130] Xia, Q., Lee, M.V., Rose, C.M., Marsh, A.J., Hubler, S.L., Wenger, C.D., Coon, J.J., *J. Am. Soc. Mass Spectrom.*, **22**, (2011), 255-264.
[131] Savitski, M.M., Nielsen, M.L., Zubarev, R.A., *Anal. Chem.*, **79**, (2007), 2296-2302.
[132] Asakawa, D., Wada, Y., *J. Phys. Chem. B*, **118**, (2014), 12318-12325.
[133] Haselmann, K.F., Budnik, B.A., Kjeldsen, F., *Eur. J. Mass Spectrom.*, **8**, (2002), 461-469.
[134] Cooper, H.J., Håkansson, K., Marshall, A.G., *Eur. J. Mass Spectrom.*, **9**, (2003), 221-222.
[135] Falth, M., Savitski, M.M., Nielsen, M.L., Kjeldsen, F., Andren, P.E., Zubarev, R.A., *Anal. Chem.*, **80**, (2008), 8089-8094.

[136] Kjeldsen, F., Zubarev, R.A., *J. Am. Chem. Soc.*, **125**, (2003), 6628-6629.
[137] Wee, S., O'Hair, R.A.J., McFadyen, W.D., *Rapid Commun. Mass Spectrom.*, **16**, (2002), 884-890.
[138] Johnson, R.S., Martin, S.A., Biemann, K., Stults, J.T., Watson, J.T., *Anal. Chem.*, **59**, (1987), 2621-2625.
[139] Li, X., Lin, C., Han, L., Costello, C.E., O'Connor, P.B., *J. Am. Soc. Mass Spectrom.*, **21**, (2010), 646-656.
[140] Laskin, J., Yang, Z., Lam, C., Chu, I.K., *Anal. Chem.*, **79**, (2007), 6607-6614.
[141] Lam, A.K.Y., Ryzhov, V., O'Hair, R.A.J., *J. Am. Soc. Mass Spectrom.*, **21**, (2010), 1296-1312.
[142] Cooper, H.J., Hudgins, R.R., Håkansson, K., Marshall, A.G., *Int. J. Mass Spectrom.*, **228**, (2003), 723-728.
[143] Fung, Y.M.E., Chan, D.T.W., *J. Am. Soc. Mass Spectrom.*, **16**, (2005), 1223-1535.
[144] Chalkley, R.J., Brinkworth, C.S., Burlingame, A.L., *J. Am. Soc. Mass Spectrom.*, **17**, (2006), 1271-1274.
[145] Rumachik, N.G., McAlister, G.C., Russel, J.D., Bailey, D.J., Wenger, C.D., Coon, J.J., *J. Am. Soc. Mass Spectrom.*, **23**, (2012), 718-727.
[146] Simon, S., Gil, A., Sodupe, M., Bertran, J., *J. Mol. Struct.: THEOCHEM*, **727**, (2005), 191-197.
[147] Gil, A., Simon, S., Rodriguez-Santiago, L., Betran, J., Sodupe, M., *J. Chem. Theory Comput.*, **3**, (2007), 2210-2220.
[148] Ryzhov, V., Lam, A.K.Y., O'Hair, R.A.J., *J. Am. Soc. Mass Spectrom.*, **20**, (2009), 985-995.
[149] Hao, G., Gross, S.S., *J. Am. Soc. Mass Spectrom.*, **17**, (2005), 1725-1730.
[150] Wee, S., O'Hair, R.A.J., McFadyen, W.D., *Int. J. Mass Spectrom.*, **234**, (2004), 101-122.
[151] Osburn, S., Steill, J.D., Oomens, J., O'Hair, R.A.J., van Stipdonk, M., Ryzhov, V., *Chem. Eur. J.*, **17**, (2011), 873-879.
[152] Sinha, R., Maitre, P., Piccirillo, S., Chiavarino, B., Crestoni, M., Fornarini, S., *Phys. Chem. Chem. Phys.*, **12**, (2010), 9794-9800.
[153] Shaffer, C.J., Marek, A., Pepin, R., Slovakova, K., Tureček, F., *J. Mass Spectrom.*, **50**, (2015), 470-475.
[154] Oomens, J., Sartakov, B.G., Meijer, G., von Helden, G., *Int. J. Mass Spectrom.*, **254**, (2006), 1-19.
[155] Kempkes, L.J.M., Martens, J., Berden, G., Oomens, J., *Rapid Commun. Mass Spectrom.* , **30**, (2016), 483-490.

Samenvatting

Op dit moment zijn er diverse technieken beschikbaar om de drie-dimensionele structuur van eiwitten te bepalen. Voorbeelden zijn NMR-spectroscopie, elektronenmicroscopie en röntgenkristallografie. Echter zijn deze methoden niet altijd geschikt, bijvoorbeeld omdat het eiwit te groot is, of omdat het moeilijk is om geschikte eiwitkristallen te groeien. Hoewel de drie-dimensionele structuur van eiwitten zeer complex kan zijn, is de primaire structuur van eiwitten relatief eenvoudig. Alle eiwitten zijn namelijk opgebouwd uit kleinere bouwstenen, de aminozuren, waarin er in de natuur precies 20 verschillende van voorkomen.

Om de aminozuurvolgorde in een eiwit te bepalen (sequencen) wordt het eiwit op specifieke, vooraf bekende plekken in de keten opgebroken. Dit kan zijn na een bepaald aminozuur door het gebruik van enzymen die het eiwit alleen op die plek in de keten op kunnen knippen (proteases). Hierdoor onstaan kortere stukjes eiwit, die peptiden worden genoemd. De volgorde van individuele aminozuren binnen deze peptides wordt bepaald door het verbreken van de peptidebinding tussen twee aminozuren in een massaspectrometer, een instrument dat de massa's van zowel de intacte moleculen als de fragmenten kan bepalen. De massa's van de resulterende fragmenten worden vervolgens afgelezen in het massa spectrum. De fragmenten die ontstaan door het eiwit specifiek te breken tussen twee aminozuren in worden sequence ionen (letterlijk volgorde ionen) genoemd. Het massaverschil tussen twee sequence ionen is gelijk aan de massa van 1 aminozuur residu. Doordat de massa's van alle 20 aminozuren bekend zijn, kan dankzij het massaverschil het bijbehorende aminozuur toegewezen worden.

Tegenwoordig zijn er speciale computeralgoritmes ontwikkeld die een massa spectrum kunnen aflezen en daarmee automatisch de aminozuur volgorde kunnen bepalen. Helaas zijn deze computer programma's nog niet waterdicht. Eiwitten kunnen namelijk op meerdere plekken breken dan zojuist is genoemd. Vaak breken ook delen van de aminozuur-zijketens af, voorbeelden hiervan zijn verlies van neutrale ammonia- en watermoleculen. Dit leidt tot een wirwar van extra pieken met een lage intensiteit in het massa spectrum, die vaak niet worden herkend door de computeralgoritmes. Doordat niet alle pieken in het massa spectrum gebruikt worden, gaat mogelijk belangrijke informatie over de eiwitstructuur verloren. Dit heeft bijvoorbeeld als gevolg dat detectie van post-translationele modificaties (PTMs) bemoeilijkt wordt.

PTMs zijn modificaties in eiwitten die plaatsvinden nadat het eiwit gesynthetiseerd is uit translatie van het RNA. Voorbeelden zijn de toevoeging van fosfaat-, methyl- en suikergroepen aan de zijketens van het eiwit. Deze modificaties kunnen spontaan optreden zoals door verlies van een ammonia of water groep, maar kunnen ook gekatalyseerd worden door enzymen, zoals het geval bij glycosylatie (toevoeging van suikergroepen). Omdat PTMs ten grondslag kunnen liggen aan diverse ziektes is het belangrijk om deze op te kunnen sporen in eiwitten.

Het doel van het onderzoek dat in dit proefschrift gedaan is, is om de chemische reacties fundamenteel te begrijpen die plaatsvinden tijdens het fragmenteren van eiwitten en peptides in de massaspectrometer. Wij denken dat hiermee uiteindelijk de computeralgoritmes verbeterd kunnen worden, zodat de structuur van eiwitten beter en sneller achterhaald kan worden en PTMs beter gevonden en geïdentificeerd kunnen worden.

Om eiwitten te fragmenteren in een massaspectrometer zijn diverse technieken beschikbaar. Elke methode breekt het eiwit op specifieke plaatsen in de eiwitketen en daarmee heeft elke techniek zijn eigen voor- en nadelen. In dit proefschrift zijn “Collision Induced Dissociation” (CID) en “Electron Transfer Dissocation” (ETD) toegepast. Tijdens CID (botsingsgeïnduceerde fragmentatie) ondergaat het eiwit botsingen met moleculen of atomen die niet reactief zijn, zoals het edelgas Helium, waarbij de botsingsenergie dissociatie van het eiwit induceert. In dit geval breekt het eiwit altijd op de peptide binding en dit leidt tot de vorming van zogenaamde b- en y-type sequence ionen (zie figuur 1-8).

Daarnaast vinden er ook veel verliezen plaats van kleine neutrale moleculen zoals ammonia, water en de groepen die als PTMs aan het eiwit zijn gekoppeld. Juist doordat deze groepen van de aminozuur residuen afbreken tijdens het sequencen van eiwitten met CID is dit niet de geschikte methode om PTMs op te sporen. Om deze reden wordt CID vaak in combinatie gebruikt met ETD, een techniek die minder gevoelig is voor het verliezen van PTMs. Tijdens ETD, wat letterlijk vertaald dissociatie door elektronoverdracht betekent, vangt het (meervoudig positief geladen) eiwit een elektron in. Vervolgens vindt een reorganisatie reactie plaats waarin het gevangen elektron verplaatst wordt en breekt het eiwit na de peptide binding. Dit leidt tot c- en z-type sequence ionen, in plaats van de b- en y-type ionen die ontstaan na CID. Bij ETD wordt er altijd één extra elektron toegevoegd aan het eiwit, waardoor het overgaat van een even-elektron systeem naar een oneven-elektron systeem, een radicaal dus. De chemische reacties van deze radicalen, geïnitieerd door de energie die vrijkomt bij de recombinatie van

positieve en negatieve lading, leidt vaak tot moeilijk interpreteerbare massaspectra. Pieken kunnen namelijk met een massa van 1 eenheid verschillen door het overspringen van een H-atoom van het z-type ion naar het c-type ion. Daarnaast zijn reactiemechanismen die plaatsvinden tijdens ETD nog onvoldoende begrepen. Onbeantwoorde vraagstukken in de literatuur zijn bijvoorbeeld de plek waar het elektron wordt ingevangen door het eiwit, wat de volgorde van processen is die plaatsvinden voordat het eiwit breekt en wat de structuren zijn van de fragmenten die door de ETD reactie ontstaan.

Hoewel massaspectrometrie toegepast wordt om de aminozuur volgorde van eiwitten te bepalen, blijft de precieze moleculaire structuur van het complete eiwit en de eiwitfragmenten onbekend. Infrarood spectroscopie is een veel gebruikte methode om de structuur van moleculen te bepalen. In een standaard spectrofotometer wordt een vloeistof of oplossing in een cuvet bestraald met infrarood licht. Het verschil tussen de invallende en uittredende lichtsterkte wordt gebruikt om de absorptie door het monster te bepalen. Elk molecuul neemt namelijk infrarood licht op van specifieke frequenties die karakteristiek zijn voor de moleculaire structuur. Maar doordat de fragmenten die gecreëerd worden in de massa spectrometer niet overgedragen kunnen worden in een spectrofotometer en omdat de concentratie van de moleculen te laag is om een verschil in absorptie te kunnen meten, is een andere techniek nodig.

Onze massaspectrometer is daarom gekoppeld aan de vrije elektronen laser FELIX, die in frequentie afstembaar infrarood licht genereert. Door een venster aan te brengen in de massaspectrometer, kunnen de fragmentionen in de ionenval met het licht van FELIX bestraald worden en kan een infrarood spectrum opgenomen worden. Maar het meten van alleen de infrarood spectra van de fragmenten is niet voldoende om de structuur te bepalen. Door theoretische infrarood spectra te berekenen van alle mogelijke structuren die de fragmenten zouden kunnen aannemen en door die vervolgens te vergelijken met het experimentele spectrum, kan heel precies worden bepaald wat de structuur is. Deze methode is zelfs zo precies dat verschillen in de vouwing van het molecuul (conformeren) herkend kunnen worden. Als het experimentele spectrum hetzelfde is als een van de berekende spectra wordt gesproken van een match, en daarmee is de structuur van het fragment bepaald. Als de structuur van het eindproduct van de chemische reactie bekend is, kan vervolgens het bijbehorende reactiemechanisme herleid worden en daarmee is het doel van dit proefschrift behaald: het dissociatie mechanisme is bekend.

In dit proefschrift zijn de structuren van de fragmenten bepaald die zijn ontstaan door CID en ETD te gebruiken op geprotoneerde peptiden. Het resultaat hiervan is dat er meer inzicht is verkregen in de verliezen van kleine neutrale moleculen tijdens CID en zijn belangrijke vraagstukken in de literatuur over het dissociatie mechanisme van ETD beantwoord. Dit proefschrift is opgedeeld in 3 delen; het eerste deel geeft een algemene inleiding op eiwit-fragmentatie en de gebruikte onderzoeksmethodes. Deel twee behandelt de structuren van fragmenten die verkregen zijn door CID. Het laatste deel beschrijft de structuren die ontrafeld zijn voor de ETD-fragmenten.

In **hoofdstuk 1** wordt uitgelegd wat eiwitten en aminozuren zijn, hoe de structuur van eiwitten bepaald kan worden door middel van massaspectrometrie en worden de methodes voor eiwitfragmentatie en hun voorgestelde chemische reactiemechanismen behandeld. Dit hoofdstuk geeft daarnaast een korte inleiding op de structuurbepaling van de eiwitfragmenten. **Hoofdstuk 2** beschrijft de gebruikte techniek om infrarood spectra te meten van ionen in de massa spectrometer, die daar slechts in zeer lage dichtheid voorkomen. Ook is hier de gebruikte experimentele opzet beschreven, hoe het laserlicht van de FELIX- en OPO-lichtbronnen gebruikt wordt, en op welke principes deze werken. Als laatste worden de theoretische methodes behandeld die gebruikt zijn voor het berekenen van de theoretische infrarood spectra.

Deel twee beschrijft de fragmentatie van aminozuren en dipeptiden met behulp van CID. Voor de bestudeerde aminozuren en dipeptiden zijn verlies van ammoniak (deamidatie) en water (dehydratie) belangrijke reactiekanalen. De gebruikte aminozuren asparagine en glutamine staan erom bekend makkelijk ammoniak te verliezen, maar er zijn veel mogelijkheden voor het reactiemechanisme waaronder dit gebeurt. Het ammoniak molecuul kan namelijk afgesplitst worden van de zijketens van deze aminozuren maar ook van de amino groep aan de N-terminus.

In **hoofdstuk 3** vangen we aan met de meest eenvoudige systemen, de deamidatiereactie van enkel de aminozuren Asparagine (Asn) en Glutamine (Gln). Door alle mogelijke opties door te rekenen is vastgesteld dat het ammoniak molecuul is afgesplitst van de zijketen in beide aminozuren. Het reactiemechanisme waaronder dit plaatsvindt is echter verschillend. In **hoofdstuk 4** zijn deze twee aminozuren gekoppeld aan een Alanine (Ala) residu en de deamidatiereacties van vier dipeptiden (AlaAsn/AlaGln/AsnAla & GlnAla) zijn bestudeerd. De invloed van de lengte van de zijketen en de aminozuurvolgorde op de deamidatiereactie is daarmee in kaart gebracht. In alle gevallen vond de

deamidatie plaats van de zijketen van asparagine of glutamine, maar wederom zijn de gevonden reactiemechanismen voor elk systeem verschillend. Daarnaast is een vergelijking gemaakt met de losse aminozuren die zijn bestudeerd in hoofdstuk 3. In **hoofdstuk 5** is de deamidatie van het dipeptide Asparagine-Valine bestudeerd en vergeleken met de deamidatie van AsnAla beschreven in hoofdstuk 4. Voor een nog gedetailleerder begrip over de deamidatiereactie zijn de resultaten verkregen met FELIX vergeleken met andere, door Amerikaanse collega's uitgevoerde experimenten, waarin de energie die nodig is voor de deamidatiereactie nauwkeurig bepaald kan worden. Daarnaast zijn kwantum-chemische berekeningen van de overgangstoestanden uitgevoerd. Dit onderzoek is gedaan in samenwerking met Georgia Boles en Peter Armentrout van de University of Utah.

Hoofdstuk 6 beschrijft de computersimulaties die zijn uitgevoerd in samenwerking met Riccardo Spezia aan de Sorbonne Universiteit in Parijs. De deamidatiereactie van twee dipeptiden (AsnAla en GlnAla) zijn gesimuleerd met de computer en de uitkomsten zijn vergeleken met de experimentele resultaten die zijn beschreven in hoofdstuk 4. Deze dynamische berekeningen op een semi-empirische potentiaal laten zien dat de deamidatiereacties vrij nauwkeurig nagebootst kunnen worden. In **hoofdstuk 7** zijn dezelfde vier dipeptiden als in hoofdstuk 4 bestudeerd maar in plaats van de deamidatiereactie is nu de dehydratiereactie onderzocht. Waterverlies kan theoretisch gezien plaatsvinden van de zijketens van de asparagine en glutamine residuen, maar ook van de carboxygroep aan de C-terminus van de dipeptiden. In dit laatste geval ontstaan er b-type ionen die een oxazolone of een diketopiperazine structuur hebben. Uit de experimenten is gebleken dat voor het dipeptide GlnAla de dehydratie plaatsvindt vanaf het glutamine-residu. In de andere gevallen geschiedt het verlies vanaf de carbonzuur groep en zijn zowel de diketopiperazine als de oxazolone structuren geobserveerd.

Deel drie beschrijft de fragmentatie van tetra- en heptapeptiden (4 en 7 aminozuurresiduen) door elektron-transfer dissociatie (ETD). De onderzochte peptiden hebben een basisch lysine residu (Lys) in de eerste positie om de vorming van c-ionen te bevorderen. In de serie van 6 heptapeptiden is het Lys-residu juist in de laatste positie opgenomen om vooral z-ionen te verkrijgen na ETD. **Hoofdstuk 8** behandelt de c-ionen. In de literatuur is een discussie gaande of deze c-type fragmenten een amide of een enol/imine structuur hebben, gebaseerd op twee verschillende onderliggende mechanismen. Door de gemeten infrarood spectra te vergelijken met berekende infrarood spectra voor zowel de amide als de enol/imine structuur blijkt dat alle gemeten c-ionen een amide structuur bezitten.

In **hoofdstuk 9** worden de structuren van de z-ionen bestudeerd. Met behulp van berekende en experimentele infrarood spectra is de plaats van het radicaal (dat wil zeggen de positie waar één waterstofatoom ontbreekt) bepaald van zes z-type ionen. In vier gevallen bleef het radicaal op de oorspronkelijke plaats, de plek waar de backbone van de peptide gesplitst is. In twee andere gevallen is waarschijnlijk een mix van twee tautomere structuren gevonden, waaruit afgeleid kan worden dat het radicaal (dus het waterstofatoom) kan migreren van de oorspronkelijke alfa-positie in de eiwitketen naar de alfa-positie in een naastliggend aminozuurresidu. **Hoofdstuk 10** beschrijft de structuur van peptiden die het aminozuur cysteïne bevatten. In plaats van de gebruikelijke z-ionen worden hier vooral w-ionen gevormd, door het verlies van een thiol radicaal van het cysteine residu dat gepaard gaat met dissociatie van de peptideketen.

Summary

A variety of methods is available to determine the three-dimensional structure of a protein. Examples are NMR-spectroscopy, electron microscopy and X-ray crystallography. However, these methods are not always useful as some proteins can be too large for NMR-spectroscopy or the growth of appropriate protein crystals can be challenging crystallization. Although the three-dimensional structure of proteins can be quite complex, it can easily be deduced to a simple primary structure: the sequence of amino acid residues, the small building blocks of every protein.

To determine the sequence of the amino acids in a protein, it is usually first broken down into smaller peptides. This is accomplished by special enzymes (proteases) that cleave the protein after a specific amino acid residue. Determination of the amino acid sequence of the resulting peptides then takes place in a mass spectrometer, an analytical instrument that is capable to measure the mass of every molecule in the instrument. This results in a mass spectrum, which gives an accurate value for the molecular weights of the peptides and their fragments and also gives an indication of their relative abundances.

The resulting fragments after specific backbone cleavages are referred to as “sequence-ions”. The mass difference between two peaks in the mass spectrum resulting from the sequence ions is equal to the mass of an amino acid residue. Since the masses of all twenty naturally occurring amino acids is known, an amino acid residue can be assigned from these mass differences. Nowadays, the assignment of amino acid sequences is done by computer algorithms which are especially developed for protein identification and sequencing. However, not every peak in the mass spectrum is recognized by these sequencing algorithms, which is for instance challenging for the identification of modified proteins. Protein fragmentation can not only occur at the backbone, but also at other positions, leading to loss of side chains or small neutral molecules such as ammonia and water. These losses lead to additional peaks in the mass spectrum and can be a cause of protein misidentification, or the assignment of the incorrect amino acid residue when not properly recognized. Especially the sequencing of post-translational modifications (PTMs) is an important challenge for protein sequencing. PTMs constitute alterations in a protein that are induced by an enzymatic reaction or that can occur spontaneously. Examples of PTMs are attachment of phosphate, methyl and carbohydrate (sugar) groups, or loss of water

and ammonia molecules from residue side chains. Because PTMs can be the underlying cause of several diseases, it is of importance to obtain a detailed characterization of their position along the protein backbone.

The general aim of the research that is performed in this thesis is to obtain a better understanding of the chemical reactions, at the level of the actual molecular structures, that take place during protein fragmentation inside mass spectrometers. With this information, sequencing algorithms can be eventually improved, so that proteins are more accurately identified and PTMs can be detected. This information can then be used for a better understanding of diseases, but also for the development of new pharmaceuticals.

Several techniques exist that are able to dissociate proteins within the mass spectrometer. Each method has its own advantages and disadvantages, as it leads to dissociation of specific bonds in the protein backbone. In this thesis Collision Induced Dissociation (CID) and Electron Transfer Dissociation (ETD) are used for protein fragmentation. During CID, a protonated peptide undergoes collisions with a neutral gas, which is usually inert and does not interfere in the reaction mechanisms, other than energizing the peptide so that the collision energy leads to cleavage of the protein, commonly at one of the amide bonds. This is the point in the backbone where two amino acid residues are linked to each other. The energy that results from the collision is redistributed over the molecule and leads to dissociation of the weakest bonds in the protein, this results in the formation of b- and y-type sequence ions. Loss of small neutral molecules such as ammonia and water is also a common dissociation pathway.

The disadvantage of collisional dissociation is that PTMs are often easily detached from the protein, making it impossible to determine the original position of the PTM in the protein. For this reason another dissociation technique has rapidly gained popularity: ETD, which is often used in combination with CID providing complementary sequence information. During ETD, an electron is captured by the multiply charged protein, which leads to dissociation of the protein at one of the backbone N-Cα bonds to produce c- and z-type sequence ions. ETD leaves PTMs mostly intact, which makes it a more appropriate method for sequencing modified proteins. However, the chemistry underlying electron transfer dissociation remains elusive. Furthermore the resulting mass spectra are more difficult to interpret as migration of an H-atom from the z-type ion to the c-type ion can occur, leading to mass peaks which are 1 mass unit higher or lower. Unanswered questions involve the position at which the electron is captured in the protein, the sequence of

events leading to protein cleavage and the molecular structure is of the resulting fragments.

Although mass spectrometry is a useful technique to determine the primary structure of a protein, the molecular structure of the complete protein and the protein fragments often remain unknown. Infrared spectroscopy is a well-known method to determine the structure of molecules. In a spectrophotometer, the absorbance of infrared light is used to determine the structure of molecules, as functional groups absorb light at specific wavelengths. This method can however not be applied to the peptide fragments, as they are created in an ion trap mass spectrometer, where they occur at extremely low densities, and it is impossible to transfer them into an photospectrometer. The low abundance precludes the observation of any absorbance of the infrared light. Therefore, the mass spectrometer is coupled to the free electron laser FELIX, which produces wavelength-tunable infrared laser light. The mass spectrometer is able to fragment the peptide ions, as well as to store the fragmented ions. Two holes are made in the trapping electrodes of the ion trap mass spectrometer, so that the light of the laser can access the ion cloud. In this way, infrared spectra can be recorded for the mass isolated ion of interest.

Measuring an infrared spectrum of a molecule alone is not enough to determine its structure. Molecular structures are derived from the IR spectra by comparing the experimental spectra with calculated spectra. Using quantum-chemical calculations, infrared spectra are calculated for structures deemed possible for the mass that is isolated and whose IR spectrum is measured. This method of ion spectroscopy allows us to determine molecular structures very precisely. It gives information beyond what a mass spectrum alone can provide and it can for instance determine the conformeric state (folding) of the molecule, as a slight change in conformation also changes the IR spectrum. In this way, molecular structures of the product ions are determined from which reaction mechanisms can be implicated. This method is used in this thesis for product ion determination and subsequently to suggest the involved reaction mechanisms for the fragmentation reactions by CID and ETD. Eventually this information can be used to improve the sequencing algorithms.

In this thesis, the fragments that resulted from CID and ETD on peptides have been studied using a combination of infrared ion spectroscopy and quantum-chemical calculations. Neutral molecule losses from amino acids and asparagine and glutamine containing dipeptides upon CID were studied. Furthermore, structures of

various types of ETD fragments have been determined in order to better understand the underlying reaction mechanisms. This thesis is divided in three parts, with the first part giving a general introduction to protein sequencing and the experimental techniques employed. Part two contains investigations into the product structures that are obtained by applying CID on a set of amino acids and dipeptides containing asparagine or glutamine. The last part describes the structures of ETD-induced fragment ions.

In **chapter 1** a brief introduction into the structure of proteins and their constituent amino acids is given. A description of methods used to determine protein structure is given in more detail focusing on protein fragmentation techniques employed in current mass spectrometry based analysis. This chapter also gives an introduction into the method of choice in this thesis, infrared ion spectroscopy in combination with quantum-chemical calculations. **Chapter 2** describes the experiment in more detail. This chapter starts with an explanation of how one can measure infrared spectra of the low-density ions in a mass spectrometer. Then the chapter continues with a description of the experimental set-up, detailing specifics of the ion trap mass spectrometer and the laser sources used, FELIX and an OPO. This chapter concludes with a short description of the quantum-chemical calculations that form an integral part of our strategies for structure determination and reaction mechanism simulations.

Part two contains investigations into the product ion structures that are obtained by applying CID on the amino acids asparagine (Asn) and glutamine (Gln) and several dipeptides containing these amino acids, coupled with an alanine (Ala) residue. Loss of an ammonia molecule (deamidation) and loss of a water molecule (dehydration) were the most prominent reaction channels observed. The amino acids Asn and Gln were specifically selected for these studies as they are well known to undergo neutral losses of NH_3 and H_2O. These losses can be hypothesized to occur from either the amino acid side chains or from the peptide backbone. Although loss of a neutral molecule may appear to be a simple reaction, the underlying mechanistics are diverse in each of the systems studied. Cyclization reactions are common but a wide range of different product structures is observed.

Chapter 3 starts with the most simple deamidation reactions, that of the amino acids Asn and Gln. In both cases we conclude that the ammonia molecule is lost from the side chain residue. In **chapter 4** these two amino acids were coupled to an alanine residue to form the dipeptides AlaAsn, AlaGln, AsnAla and GlnAla. The product structures after loss of an ammonia molecule from these dipeptides were

determined. The influence of the sidechain length and the amino acid sequence on the deamiation reaction were studied. Although the ammonia molecule was lost from the Asn or Gln side chain in all cases, different product structures were established for each system. The deamidation of dipeptide Asparagine-Valine (AsnVal) is described in **Chapter 5** and compared with the deamidation mechanism of AsnAla. For a more detailed understanding of the underlying mechanisms, the FELIX results were compared with experiments and calculations performed by Georgia Boles and Peter Armentrout from the University of Utah. A combination of infrared spectroscopy, threshold-CID experiments (in which the reaction energies were accurately determined) and transition-state calculations was used. The AsnAla and GlnAla peptides that were studied in chapter 4 have been used as a benchmark for computational simulations of the deamidation reaction. The outcome of these simulations is described in **Chapter 6** and is compared with the experimental results from chapter 4. It turned out that the simulated reaction pathways are largely in accordance with the ones found in the experiment. This research has been carried out at the Sorbonne University in Paris in collaboration with Riccardo Spezia. In the last chapter of the second part, **Chapter 7**, the dehydration reaction for the same four peptides as in chapter 4 was studied. The dipeptide GlnAla loses a water molecule from the Gln sidechain, whereas for the other three dipeptides the water molecule is lost from the backbone, leading to b-type ions that possess an oxazolone and a diketopiperazine structure.

The last part of this thesis describes several investigations of the structures of fragment ions that resulted after applying ETD on a set of tetra- and heptapeptides (four and seven amino acid residues). A basic lysine residue (Lys) is introduced at the first position in the peptide to enhance the formation of c-type ions. To trigger the formation of C-terminal w- and z-type ions, a Lys residue is selected as the last residue of the peptide. The structure determination of c-ions is described in **Chapter 8**. In the literature, c-type ions have been suggested to possess either an amide or an enol/imine structure. Our measurements for a set of 9 c-ions showed that an amide structure is formed in all cases. The structure of the complementary z-type ions is investigated in **Chapter 9**. Using a combination of theoretical calculations and experiments, the location of the radical in these z-ions has been determined. Of the six z-type ions that were studied, four cases showed no trace of radical migration, i.e. the radical remains on the position of the cleavage of the backbone, the Cα-atom of the cleaved residue. For the other two cases, a combination of two structures, no radical migration and radical migration to the adjacent amino acid residue, appears to be most likely. In the final **Chapter 10** of

this thesis, the structure of w-type ions has been spectroscopically investigated. These w-type ions result from side-chain loss concomitant with ETD backbone cleavage, and is for instance observed for cysteine residues, that are prone to losing a thiol radical from their side chain upon ETD activation.

Publications

Publications of the work described in this thesis:

- **L.J.M. Kempkes**, J. Martens, J. Grzetic, G. Berden and J. Oomens
  Deamidation reactions of protonated asparagine and glutamine investigated by ion spectroscopy. *Rapid Communications in Mass Spectrometry, Volume 30, Issue 4, pp 483-490 (2016).* DOI: 10.1002/rcm.7464

- **L.J.M. Kempkes**, J. Martens, J. Grzetic, G. Berden and J. Oomens
  Deamidation reactions of asparagine- and glutamine- containing dipeptides investigated by ion spectroscopy.
  *Journal of The American Society for Mass Spectrometry, Volume 27, Issue 11, pp 1855-1869 (2016).*DOI: 10.1007/s13361-016-1462-5

- **L.J.M. Kempkes**, J. Martens, G. Berden and J. Oomens
  Dehydration reactions of pronated dipeptides containing asparagine or glutamine investigated by infrared ion spectroscopy.
  International Journal of Mass Spectrometry, Volume 429, pp 90-100 (2018).
  DOI: 10.1016/j.ijms.2017.06.004

- **L.J.M. Kempkes,** G.C. Boles, J. Martens, G. Berden, P.B. Armentrout and J. Oomens,
  Deamidation of Protonated Asparagine-Valine Investigated by a Combined Spectroscopic, Guided Ion Beam, and Theoretical Study.
  The Journal of Physical Chemistry A, Volume 122, Issue 9, pp 2424-2436 (2018).
  DOI: 10.1021/acs.jpca.7b12348

- **L.J.M. Kempkes,** J. Martens, G. Berden and J. Oomens
  w-Type ions formed by Electron Transfer Dissociation of Cys-containing peptides investigated by infrared ion spectroscopy.
  Journal of Mass Spectrometry, Volume 53, pp 1207-1213 (2018).
  DOI: 10.1002/jms.4298

- **L.J.M. Kempkes**, J. Martens, G. Berden and J. Oomens
  Spectroscopic characterization of an extensive set of c-type peptide fragment ions formed by electron transfer dissociation suggest exclusive formation of amide isomers. *Journal of Physical Chemistry Letters, Volume 9, pp 6404-6411 (2018).* DOI: 10.1021/acs.jpclett.8b02850

- **L.J.M. Kempkes**, J. Martens, G. Berden, K.J. Houthuijs and J. Oomens
  Investigation of the position of the radical in z_3-ions resulting from electron transfer dissociation using Infrared Ion Spectroscopy. *Faraday Discussions, In Press.* DOI: 10.1039/C8FD00202A

Other publications:

- **L.J.M. Kempkes** and W.J.P. van Enckevort
  On the concomitant crystallization of amino acid crystals upon dissolution of some amino acid salt crystals. *CrystEngComm, Volume 17, Issue 6, pp 1483-1490 (2015).*
  DOI: 10.1039/C4CE02070G

- G.C. Boles, **L.J.M. Kempkes**, J. Martens, G. Berden, J. Oomens, P.B. Armentrout
  Ion spectroscopy and guided ion beam studies of protonated asparaginyl-threonine decomposition: influence of a hydroxyl containing C-terminal residue on the deamidation process.
  Manuscript in preparation

- **L.J.M. Kempkes,** G.C. Boles, J. Martens, G. Berden, P.B. Armentrout and J. Oomens
  The influence of a hydroxyl-group on the deamidation reaction of protonated asparagine-serine investigated by a combined spectroscopic, guided ion beam and theoretical study.
  Manuscript in preparation

- T. Geurts, **L.J.M. Kempkes**, J. Martens, G. Berden, J. Oomens
  Dehydration reactions of pronated dipeptides containing asparagine: the influence of the alcohol group containing side chains investigated by infrared ion spectroscopy.
  Manuscript in preparation